THE RELATIONSHIP BETWEEN COMPETENCE AND PERFORMANCE IN EARLY DEVELOPMENT IN CHILDREN WITH DOWN'S SYNDROME

Louise A. Duffy

Thesis submitted for
Degree of Ph D

University of Edinburgh

December 1990

DECLARATION

I declare this thesis has been composed by myself and that the work contained therein is my own.

Louise A. Duffy

ACKNOWLEDGEMENTS

Firstly, I must acknowledge, with very many thanks the constant help, support, guidance and patience of my supervisor Jennifer Wishart without whom this thesis would not have been possible, either in its inception or in its realisation. I am also extremely indebted to Sue Widdecombe who has contributed in many ways to this piece of work and who has put up with me as a (very trying) flatmate over the last two years. A great many thanks too, to Bob for helping to keep things in perspective and to Henry for his humour and patience throughout a difficult time.

Many, many thanks to all of the children who participated in the studies presented here and to the parents who, time and time again, brought the younger ones into the Psychology Department. I am also very grateful to the head teachers and staff at Prospect Bank, St. Crispins and Kingsinch schools, and to the staff at the Psychology Department and Colinton nursery schools. Thanks also to the Scottish Down's Syndrome Association for all their support.

This research was supported by ESRC linked Studentship Award No. C0042872500B and by MRC Project No. G8703875N.

ABSTRACT

This thesis investigated the effects of performance deficits on the development of cognitive competence in children with Down's Syndrome (DS). Fifty-one DS subjects ranging in age between 3 months and 16 years participated in 5 cross-sectional and 2 longitudinal studies. Performance was studied both in a standardised assessment situation and in response to a set of discrimination learning tasks.

Each of the 4 assessment studies revealed a substantial differential between levels of performance and true competence in this group of DS subjects. Although overall levels of 'IQ' found were similar to those reported in previous studies of cognitive ability in DS children, item-item analyses and analyses of test behaviour revealed a strong tendency to underperform. Task avoidance was frequent, resulting in very poor levels of test-retest reliability. A pattern emerged from the longitudinal data suggesting that there may be a causal relationship between this test-retest unreliability and the apparent loss of skills at later age levels. It was found that children often withheld demonstration of optimal levels of performance on tasks with which they had encountered difficulty at younger age levels.

The implications of the results from the assessment studies are three-fold: firstly, the repeated evidence of unreliable performance and of developmental instability suggests that psychometric methods of cognitive assessment should be used with great caution with this population of mentally handicapped children; secondly, and relatedly, it suggests that theories of development in DS based on outcome measures obtained from such tests may be inaccurate; thirdly, the frequency with which children were seen to 'lose' skills suggests that teaching effort should be focussed not only on encouraging the development of new skills but also on ensuring the consolidation of these skills once acquired.

Three discrimination learning studies investigated whether an errorless learning procedure might enhance motivation to perform to full competence and increase performance reliability. Results from the first two studies were highly encouraging: with this technique DS subjects quickly learned the target discrimination and this skill was reliably demonstrated over 4 testing sessions. Results from a third study indicated, however, that the technique was of little value in teaching learning skills per se: learning did not transfer to a second, very similar discrimination task. It was concluded that while errorless learning procedures have a useful role to play in teaching DS children that they can learn, they may be most effective when used in conjunction with conventional trial-and-error teaching methods.

CONTENTS

CHAPTER 1

INTRODUCTION

This thesis will present a study of early cognitive development in children with Down's Syndrome (DS), placing particular emphasis on the relationship between performance and competence in the development of learning difficulties in this population. Although the focus will be on Down's Syndrome, it is hoped that findings may be of relevance to the field of mental handicap in general.

The study of development in handicap: practical considerations

It is encouraging to note that the study of development in handicap would at last appear to be receiving the research attention it deserves. At the recent European Conference on Developmental Psychology (August 1990) a substantial number of presentations concerned a variety of handicapping conditions. Given the sheer size of the population of individuals with handicaps this recent increase in attention could hardly be seen to be unwarranted. Warnock (1978) concluded that "up to one in five children at some time in their school career will require some form of special educational provision". Wynn and Wynn (1979) reported that about 1% of the U.K. population are so severely handicapped that they will require help throughout their lives with many or all basic skills. A further 10% are so severely handicapped that it is unlikely that they will find more than the humblest sort of employment. Wynn and Wynn estimated that in Britain 130,000 children under the age of 15 are very severely handicapped, and that a further 30,000 handicapped children are being born every year. In the U.S.A., Babson and Benson (1971) estimate that 50,000 babies a year are born with severe malformations, and an additional 300,000 have some degree of learning difficulty, ranging from mild to severe.

Clearly, not only do the handicapped represent one of the most disadvantaged sections of the population, they also represent a significant proportion of that population. In their own right they have a clear case for claiming a proportionate amount of the resources available for research. It could be further argued, that given the severity of their disadvantage there is indeed a case for claiming a greater than proportionate amount of these resources.

Today, because of the nature of recent reforms in government policy towards handicapped groups, the claim to these resources could be seen to be even more pressing than ever before. One such series of reforms relates to the new impetus to care for the mentally handicapped in the community. Following the launch of the DHSS 'Care in the Community' initiative (1983) there has been a shift in policy towards the more rapid closure of hospitals for the mentally handicapped and an associated switch of resources from health boards to local authorities. In conjunction with these changes, as a result of the publication of the Warnock Report (1978), there has also been a move towards integrating mentally handicapped children into mainstream schools.

To ensure the successful implementation of the above reforms it would seem vital that we continue to increase our knowledge of the specific needs of individuals with mentally handicapping conditions. In particular, there can be little doubt that educational integration can only be effective if those who work within the mainstream education system are adequately informed about the development of children with mental handicap.

Practical problems in working with handicapped populations

Although there is a pressing need for more knowledge of the development of children with mental handicap, it is important to acknowledge, at this point, that there are a number of practical problems in carrying out research with this population. Before discussing these, it is perhaps worth noting that there are problems inherent in all developmental and other social science research dependent on the availability and cooperation of subjects. Epidemiologists have long been concerned with the effects of subject loss and unresponsiveness on the outcome of both experimental and non-experimental investigations (e.g., Cox, Rutter, Yule & Quinton, 1977; McMahon & Pugh,1970). Several studies have drawn attention to the importance of this problem in infant research. It has been estimated that up to 30% of infants who participate in experiments within the first half year of life are discarded as subjects because they fail to provide valid data (Lewis and Johnson 1971). Investigation of those infants excluded from data analysis on the grounds of uncooperativeness or unresponsiveness suggest that there are in fact differences between those infants who complete tests and those who do not (Lewis and Johnson 1971; DeLoache, Rissman and Cohen 1978). The results of a recent study by Wachs and Smitherman (1985) lend further support to the suggestion that incomplete data or subject loss are not randomly determined and may bias findings.

In the field of mental handicap research a number of additional problems compound those outlined above. Until the 1960s, a major confounding factor in the study of handicapped populations was the institution in which the children were living. As a consequence of a number of studies demonstrating the deleterious effects of institutionalisation on development (Tizard 1960, Stedman and Eichorn 1964, Bayley et al 1966), studies comparing handicapped institutionalised children with non-handicapped children living in their parental home were heavily criticised on the grounds that it was impossible to determine whether results reflected genuine differences in potential between the two populations or simply the effects of institutionalisation. With the move towards integrating handicapped sections of the population into society, this has become less of a problem for researchers; handicapped children are now being brought up in an environment similar to that of their non-handicapped peers.

However there are still a number of difficulties in ensuring that mentally handicapped children available for participation in research studies are representative. Many forms of handicap are associated with additional health difficulties. For example, in a study of 480 children with disabilities of varying aetiologies, it was found that over half of the sample had been admitted to hospital at least once in the previous year (Baldwin 1985). The implications of these findings are twofold: firstly, only the healthiest children may be available for inclusion in research samples and secondly, there may be a considerable amount of subject loss due to illness among those who do participate. The effects of long periods of illness and/or hospitalisation on development must also be taken into account.

The availability of subjects for psychological research is also dependent on parents volunteering their child, or at the very least giving parental consent to their being included in a particular research study. The 'costs' of having a handicapped child, in terms both of time and economics, may to some extent affect these decisions. Sample bias may arise also from differences between individual parents' inclinations to allow their children to take part in investigations. Implicit in the decision to volunteer a child for handicap research is an acceptance - which cannot always be assumed - of that child's handicap.

The study of development in handicap: its relevance to mainstream developmental psychology

Despite these considerable practical difficulties, it still remains crucial that we continue to increase our knowledge of development in handicapped children. In addition to the obvious value to the handicapped themselves, findings from handicap research have a large contribution to make to developmental psychology as a whole. It is essential that any attempt to explain the processes which contribute to successful, adaptive behaviour should also be able to explain how and why it is that these processes are interrupted in abnormal development (Cichetti and Sroufe 1978; Cichetti and Pogge-Hesse 1982; Cichetti and Shneider-Rosen 1983). This approach is not new; historically, scientists in a number of disciplines have pointed to the need for looking at the relationship between normal and abnormal patterns of development.

Comparisons of atypical and normal populations increase our understanding of how specific abnormalities can affect development. They can also help identify factors influencing development which are less obviously manifested in the development of non-handicapped children. Studies have shown that children with visual, hearing and motor handicaps can achieve developmental skills very similar to those acquired by non-handicapped children. This would seem to indicate that sight, hearing and mobility are not prerequisites for the development of specific abilities. These disabilities may, however, influence the way in which specific skills are acquired. The development of reaching in blind children, for example, has been observed to follow a different 'route' from that seen in sighted children, highlighting the role vision plays in making the association between the tactual and auditory properties of an object in sighted children. Blind children do come to make this association, but very much later, indicating that it is possible to integrate information in different ways to reach similar levels of understanding (Fraiberg 1977).

Results from studies of the development of children with mental handicap can similarly increase our understanding of the relationship between chronological age and mental age in the development of certain abilities. Because chronological and mental age are very closely linked in the normal population, it is difficult to tease the two apart in studies using only non-handicapped children. Consequently it is frequently only possible to speculate on developmental links between domains such as cognition and emotion. By investigating mentally handicapped children, an atypical population in which mental and chronological ages are not so closely interrelated, less speculative conclusions can be drawn.

Why study Down's Syndrome rather than other forms of mental handicap?

There are a number of advantages to focussing on DS in any research on early development in children with a mental handicap. DS is a chromosomal abnormality which can be identified at birth and affects a large number of children - approximately 1 in every 660 children born. Whereas other forms of mental handicap may take as long as two or even five years in some cases to diagnose, the availability at such an early stage in development of such a large number of children, guaranteed by their condition to demonstrate some degree of mental handicap, presents a vast resource for studying early development in the mentally handicapped child.

The fact that the syndrome contains within it three differing types - trisomy 21, translocation and mosaicism - is an added advantage. Mosaicism, for example, is believed by some to have a more favourable developmental prognosis than the other two forms of DS. Although there is still considerable disagreement as to whether mosaic DS children, having only a proportion of trisomic cells, are or are not generally less retarded than those with full trisomy 21, what *is* known is that some of the least retarded cases of DS are mosaic (Hamerton, Gianelli and Polani 1963, Fishler 1975). In general, however, there is substantial variation in intellectual ability among individuals with DS (see Gibson 1978), a prerequisite for any research into development and developmental processes.

The size of the population of individuals with DS would in itself seem to warrant more research than has been done to date. DS accounts for one third of the population of the severely mentally retarded. This incidence rate - which is universal - seems to have remained relatively constant over time. Prevalence, however, has increased greatly, both through the reduction of mortality rates and the extension of life expectancy.

Presently-available screening methods appear to have had disappointingly little impact on overall incidence rates. Despite recent developments based on a composite risk factor derived from serum concentrations of alpha fetoprotein, unconjugated oestriol, human chorionic gonadotrophin and maternal age (Wald et al 1988), problems inherent in this strategy make it unlikely that it will be widely adopted in the near future. Although it is claimed that this screening procedure could identify as many as 60% of all affected pregnancies, as Donnai and Andrews (1988) point out, it would introduce the problem of having to discriminate between those women with low composite risks and those who would normally have been offered amniocentesis on the basis of age alone. Current practice uses age as the basis for screening but this means that only about 30% of all DS pregnancies are potentially detectable; two thirds of DS babies are born to mothers under 35, the usual cut-off point for screening. In fact fewer than half of those women entitled to screening on age grounds actually present for amniocentesis with the result that the detection figure stands at about 15% (Wald et al 1988). Although screening has reduced the number of DS births born to older women, a reported drop in fertility in mothers over 35 (i.e. in the 'at risk' category) in conjunction with stable incidence rates may imply that there is an increase in the number of younger mothers giving birth to DS children (Mikkelsen 1977; Evans et al 1978).

As yet, there is no means of identifying which parents are most likely to be at risk of having a child with DS. There are some cases in which the rarer forms of the syndrome - translocation and mosaicism - can be inherited but as these forms of DS only account respectively for 5-6% and 1-2% of the total population with the syndrome, this would have little impact on the overall incidence figures. Down's Syndrome is therefore a condition likely to continue to be with us, certainly in the foreseeable future and as such it is important to maintain our efforts to investigate its effects on development.

Problems specific to DS

At this point it is worth discussing more specifically the problems encountered in carrying out research on development in this particular population. It is important to note, however, that the majority of these problems are health-related and are therefore to a large degree becoming surmountable. Over the past 20 years dramatic improvements in both life expectancy and the quality of life have been recorded for DS children. Whereas in 1958, 70% failed to survive the first ten years with 63% dying in the first year and 30% before four weeks of age (Carter 1958), these mortality rates have now been reduced by 50-60% (Thase 1982).

The most common cause of infant mortality in DS is still respiratory infection, but treatment with antibiotics and better general medical care has dramatically reduced the number of deaths from pneumonia. The second major cause of death in DS is heart disease; approximately one in two DS children are born with a congenital heart defect (Greenwood and Nadas 1976, Rowe and Uchida (1960). This contrasts with a reported incidence of 0.7% to 1% in the population as a whole (Richards et al 1955; Mitchell et al 1971; Keith et al 1978). Medical advances in treatment, in surgery, and in advice available for parents on day to day management of a child with heart disease have led to vast improvements in the survival rate and quality of life of the DS child born with a cardiac defect. Although there is still a high mortality rate in DS infants and children who need to undergo cardiac surgery (approximately 38%), Katlick et al (1977) found no obvious differences between these figures and the figures from comparable cardiac surgery on non-DS children. This runs contrary to the previous supposition that DS added to the mortality and morbidity of cardiac surgery.

These medical advances have not only led to an increase in the numbers of DS children potentially available for research, but also offer a greater opportunity to study their development proceeding more 'normally', without the intrusion of additional health difficulties.

Another practical problem which is being both recognised and overcome is that there is a high rate of visual and hearing problems associated with the syndrome. Visual defects have been reported to affect approximately 67% of individuals with DS (Gardner 1967). Refractive errors, short and long sightedness, and astigmatism are common (Gardner 1967, Fanning 1971) but can all be corrected with spectacles.

Hearing loss is also frequently reported in children with DS. Rigrodsky et al (1961) reported 60% of those in their study to have hearing deficits, mainly in the moderate to slight range; in a study of 24 infants Cunningham and McArthur (1981) reported that 80-85% had moderate to severe hearing losses; in a group of 51 Down's adults Keiser et al (1981) found that 74% had some degree of hearing impairment; Nolan et al (1980) reported an incidence of 69%. The majority of problems arise from otitis media, a dysfunction of the middle ear. Although the immediate results of surgery for this condition (including the insertion of grommets) are often positive, the long term results are so poor that this means of treatment has largely been abandoned (Davies 1985). However hearing aids are becoming increasingly widely used with generally favourable results.

Previously the developmental effects of auditory and visual defects were mistakenly identified as the result of mental deficiency. It has however been shown that diagnosis and treatment of such conditions can lead to significant improvement in many aspects of development (see e.g. Davies 1985).

Clearly, the improvements outlined above in relation to the health of children with DS have to a large degree removed what was previously a major obstacle to research with this population of children. It is important to recognise, however, that several additional factors associated with this condition still present problems in ensuring the representativeness of DS children in relation to the wider population of children with mental handicap. Over and above the prevalence of respiratory, cardiac, visual and hearing problems, DS children almost without exception are hypotonic at birth, i.e. have poor muscle tone; this may have effects on their early motor development not present in other mentally handicapped groups (Cowie 1970). In addition, the consistent finding that DS females demonstrate higher levels of ability than males (Brousseau and Brainerd 1928; Sternlicht and Wanderer 1962; Clements et al 1976; La Veck and La Veck 1977) reduces their representativeness of the handicapped population in general in which there appears to be little difference in ability between the sexes.

Theoretical considerations in studying development in DS children

Despite the alleviation of many of the practical difficulties discussed above, there are still a number of theoretical considerations which potentially provide obstacles to researchers studying development in DS children. Until the 1970s it was generally believed that, because DS children demonstrate identifiable physiological, biochemical and/or anatomical deficits contributing to their handicapping conditions the developmental process in DS should be conceptualised as being qualitatively different from that of MA-matched handicapped children showing no obvious organic impairment. Edward Zigler, a major proponent of this viewpoint, suggested that organically handicapped groups may follow a quite different set of developmental pathways to non-handicapped persons and may manifest a kind of cognitive functioning that differs qualitatively rather than quantitatively from normal developmental processes. A fuller discussion of this difference/delay debate can be found in Chapter 3.

More recently a group of researchers in the USA have advocated the application of a more 'liberal' developmental perspective to individuals with DS and to other groups of organically handicapped persons. With this shift in perspective a considerable amount of research attention is being paid to highlighting the similarities in the developmental processes in organically handicapped and normally developing children.

The application of this perspective has given rise to two major hypotheses on development in organically handicapped groups such as DS. The first - *the similar sequence hypothesis* - is based upon the classical developmental principle of directed change. The basic tenet of this principle is that there are regular and invariant sequences of development all leading towards a clear endpoint. Piaget, a major proponent of this position, hypothesised that the child traverses cognitive stages in invariant order, beginning with sensorimotor modes of thinking and ending with the acquisition of formal operational thought. Application of this principle in research in handicap has resulted in the search for evidence of similar sequences of development in populations such as DS.

Piaget also held that the child's progression through these stages of development was characterised by both qualitative and quantitative changes. Language acquisition, for example, involves not only such quantitative achievements as the increasing size of the child's vocabulary but also a qualitative restructuring of internal mental schemes. Whereas the sensorimotor (prelinguistic) infant has only action schemes, the preoperational child can represent events in language and other symbol systems.

The second and complementary *similar structure hypothesis* on development in mental handicap draws upon the principle of qualitative change. This hypothesis predicts that across all areas of cognitive functioning handicapped children should perform equally well from task to task - in the same way that normally developing children do.

Both hypotheses encompass the orthogenetic principle (Werner 1948, 1957) which states that development proceeds from a state of relative globality and lack of differentiation to a state of increasing differentiation, articulation and hierarchic integration.

The work of Piaget and Werner largely focussed on children's achievements in the cognitive and linguistic domains. Although the basic principles outlined by both researchers still form the backbone of more contemporary expansions of developmental theory, more emphasis is now placed on non-cognitive domains such as social, emotional, motivational and personality development. In short, it is now generally held that children are more than thinking organisms and that a complete understanding of the child cannot be achieved through examination of the development of cognitive and linguistic capacities alone. This 'expanded' theory of development has also been applied to the study of development in mentally handicapped populations.

It is also now widely accepted in mainstream developmental psychology that the process of development is most profitably studied as an integrated whole rather than as a sum of distinct and separate domains. It is accepted, for example, that failure to take into account the influence of emotional factors on the development of cognition can only result in an incomplete and inadequate picture of the complexity of the developmental process. Current developmental theory as a result has shifted emphasis away from the study of isolated fragments of social, emotional and cognitive behavioral systems, and now focusses on the interrelationships between these different systems. The following section provides a brief outline of some of the research on development in DS which has adopted this perspective.

Applying the 'expanded' developmental perspective to children with DS

Affect and temperament: A number of studies of the emergence of affective responses in DS infants have suggested that development in this domain is closely related to cognitive development. Strong correlations have been found to exist between level of cognitive development and the level of complexity of stimuli which make DS infants laugh, for example (Cicchetti and Sroufe 1976). In a later study Cicchetti and Serafica (1981) found that levels of positive affect were also related to the intensity of negative reactions to the visual cliff situation. Although infants with DS show later onset of both crying and laughter than non-handicapped infants of similar ages, studies of their responses to items designed to elicit smiling and laughter or fear and distress reactions have found that the development of positive and negative affect proceeds through a similar series of stages to that seen in their non-handicapped peers.

Such findings, it is claimed, point to the similarity of structure and of sequence in the development of affect in DS and non-handicapped children. Differences have been found between the two populations, however, in terms of the intensity of affect expressions. Although, overall, studies with both infants and toddlers have produced equivocal results, several have found children with DS to demonstrate less positive and negative affect than normally developing children of the same chronological age (Rothbart and Hanson 1983) and developmental level (Gunn et al 1981; Marcovitch et al 1986). Parent report measures have also suggested differences between DS and non-handicapped children along the dimensions of threshold attention and threshold level. The persistence of such differences across age groups (i.e infants and toddlers), it has been proposed, may reflect temperament differences between these two populations of children (Cicchetti and Ganiban 1990).

Similar findings have also emerged from comparisons of DS and non-handicapped infants with regard to the qualitative and quantitative aspects of attachment and stranger fear responses. The majority of children with DS form secure attachment relationships with their caregivers (Thompson et al 1985), and although they develop at a slower pace, it has been suggested that the attachment, affiliation and fear/wariness systems of infants with DS are similarly organised to those in normal infants (Cicchetti and Serafica 1981). Again, however, infants with DS have been shown to demonstrate less intense separation distress, longer response latencies, briefer recoveries and a smaller range of affect lability in their responses in comparison to normally developing infants of the same MA or CA (Thompson et al 1985).

A number of biological studies have produced findings which may in part explain these differences found between DS and normally developing children in relation to temperament. Studies of neurotransmitters have pointed toward decreased functioning of both the parasympathetic and sympathetic branches of the peripheral nervous system, and decreases in serotonergic activity thought to affect consciousness (Weinshilbaum et al 1971; Keele et al 1969; Casanova et al 1985; Yates et al 1980; Scott et al 1983). These difference may predispose infants with DS to be more passive or less reactive than other children of similar ages.

In addition, neuroanatomical studies have found that DS affects the maturation of the brain (Becker et al 1986; Purpura 1975; Takashima et al 1981). Investigations of the growth of neuronal networks in the brains of infants and children with DS have suggested that after a short period of growth within the first few months of life, the maturation of the brain seems to decrease relative to that of the normally developing brain. It has been speculated that these restrictions may place restraints on the rate of development and the onset of transition and changes in temperament characteristics. Delayed neurological development, it has been suggested, might also affect the development of forebrain inhibitory tracts which are important for self-regulation of both arousal and attention.

Loveland (1987) has speculated that poorly developed inhibitory mechanisms may also explain the differences in attention level and information-processing abilities which have been found to exist between DS and non-handicapped children. Studies comparing the two populations have noted that youngsters with DS generally have long attention spans and appear to be less distractible than their normally developing peers (Miranda and Fantz 1973,1974; Gibson 1978; Cohen 1981; Loveland 1987). DS children have also been noted to be more visually attentive in their interactions than normally developing infants (MacTurk et al 1985; Vietze et al 1983). If such inhibitory mechanisms are impaired, the DS child's ability to disengage from a task or to redirect attention may suffer simply because the physical ability to self-regulate is impaired.

Studies that have used selective attention as a measure of information processing have shown delays in development when tasks have required higher level processing. In both visual and auditory domains, investigations of this capacity have concluded that although individuals with DS possess the basic ability to perceive information accurately, they appear to be delayed in their ability to utilise and interpret their experiences to the same extent as normally developing infants (Fantz, Fagan and Miranda 1975; Glenn, Cunningham and Joyce 1981).

Play and communication: Slower information processing capacities may explain the delayed pace of play development also found in children with DS. As in the domains of affect and attachment, DS development in this area follows a similar sequence to that seen in normally developing infants. Pre-symbolic play is characterised by a progression from predominantly visual-tactual object exploration and manipulation to the manipulation of objects in relational and combinatorial ways. The symbolic play of children with DS, although emerging at a delayed pace, progresses through the same developmental sequences of decentration, decontextualisation and integration in object and social play that characterise the play development of non-handicapped children in early childhood (Beeghly and Cicchetti 1987; Hill and McCune-Nicolich 1981; Motti et al 1983). Levels of sophistication of play have also been found to correlate with levels of overall cognitive development (Beeghly et al 1989). The same researchers have also noted, however, that one important distinguishing characteristic of the play of children with DS is that although it does become increasingly more abstract, it is still more concrete than that of their normally developing peers.

Beeghly et al also found dimensions of social play to correlate significantly with levels of play maturity and cognitive development. Differences were found between DS and MA-matched non-handicapped children, however, in the extent to which they would engage in social play. DS children spent significantly less time engaging with their peers and were rated as being less responsive and initiating than their cognitively matched controls during social interaction.

In spite of this difference, DS children have been found to be more advanced in the pragmatic (social/communicative) aspects of language acquisition than in the more structural aspects of linguistic competence (i.e vocabulary and syntax). Indeed the major deficit most often found in the DS population is in language functioning. Despite similarities in structure and sequence for most aspects of language, results of many studies document that individuals with DS show an increasing linguistic deficit in relation to their non-verbal cognitive abilities with increasing chronological age (Miller, 1991 in press). However, Beeghly and Cicchetti (1987) found children with DS to perform significantly better than linguistically matched non-handicapped controls (but not MA-matched controls) when measures of communication and pragmatic development were considered, with longer sequences of on-topic turns and conversationally relevant turns, a greater diversity of speech acts, and more mature turn-taking skills.

Cicchetti and Ganiban (1990) have proposed that this asynchrony between the pragmatic and structural aspects of language may reflect the impact of interactions between children with DS and their caregivers rather than significant differences in the organisation of behavioral systems. Cardoso-Martins and Mervis (1985) found that mothers of prelinguistic children with DS provide lesser amounts of labelling and deictic information than do mothers of non-handicapped children. It was suggested that this may result from reduced maternal expectations concerning their children's ability to learn language. More recently Mervis (1991, in press) has considered the issue of directiveness as another possible contributory factor to the linguistic deficits of children with DS. Here it is suggested that because these children respond to their environment with dampened reactivity, caretakers may be more directive in their exchanges with their children. The low degree of contingency in such exchanges may mean, that children with DS do not learn as many words as children who experience a large amount of contingent exchanges with their parents.

In summary, the application of an expanded developmental perspective to the study of DS has resulted in a number of findings which support the 'similar sequence' hypothesis of development in this population of handicapped children. In the domains of affect, attachment, play and language, for example, children with DS have been shown to progress through similar series of stages to those traversed by their non-handicapped peers. However, a number of differences in developmental structure have also been noted to exist between the two populations. In general these differences are believed to result from differences in temperament, attention and information processing capacities which can be explained with reference to biological and neuroanatomical deficits. In addition, in relation to language, it has been suggested that the structural asynchronies found may result from an interaction between such biologically determined deficits and the linguistic environment provided for children with DS.

It remains uncertain, however, whether any explanation relating deficits inherent in the condition of DS to the wider experience of the child resulting from such deficits can also be applied to explain the delays and deficits in cognitive ability which have been so widely documented. Researchers have sought explanations in the wider experience of the child for the developmental processes found in language - where findings do not fit with attempts to impose a 'normal' structure, nor with biological findings which might indicate lack of structure. Intent on applying a developmental perspective which focusses only on sequences and structures, these researchers may not in fact be addressing the whole child to any significantly greater extent than did their predecessors - the classical developmental theorists who sought to describe 'stages' in sensory-motor or cognitive development.

Indeed it could be argued that, despite claims that the new approach focuses on the whole child, many of the non-cognitive areas which have been studied from this perspective have simply been studied for the purpose of investigating whether or not these areas of development are interrelated in the same way as they appear to be interrelated in normal development. There appear to have been few attempts to study the nature of such interactions. Although we now know that there are correlations between levels of cognitive development in individual children with DS and the stages of development reached by those same individuals in affect, attachment and play, we still remain uninformed with regard to the way in which factors related to these and other areas of function may influence the DS child's experience of learning situations. Similarly, while it has been shown that low levels of arousal and deficits in attention and information processing may affect both cognitive and non-cognitive capacities, we do not know whether the extent of their contribution is similar across all domains of development. Can it be assumed that because cognitive development has been found to correlate neatly with developments in areas such as affect, attachment etc, these deficits in non-cognitive domains are in themselves sufficient to explain the cognitive deficits in DS ?

The very methods by which this group of researchers have attempted to demonstrate interrelationships between cognitive and non-cognitive domains seem to fail to acknowledge that performance in cognitive contexts is heavily influenced by non-cognitive factors. A number of researchers in mainstream developmental psychology have pointed to the difficulties inherent in cognitive testing because of the extent to which motivational factors can affect a child's performance (McCall 1981; Scarr-Salpatek 1976; Scarr 1981; Hrncir et al 1985). Despite this, the application of the developmental perspective to the study of DS has largely resulted in assessment of the cognitive abilities of children with psychometric instruments such as the Bayley Scales and the Piagetian-based Uzigiris and Hunt Scales. Tests of this nature can give no indication of the extent to which children's performance in assessment situations actually reflects their true levels of functioning. It will be the contention of this thesis that any perspective which neglects the role of motivational factors in development cannot attempt to represent accurately the true nature of development in DS.

Mental handicap and motivation

In a recent review of the role of motivational factors in the functioning of mentally handicapped individuals, Zigler and his colleagues pointed to the need to translate research findings from this area of study into assessment techniques (Merighi, Edison and Zigler 1990). A constant dimension characterising descriptions of DS is that of a lack of belief in personal efficacy. Strongly inclined to expect failure, the retarded child is more responsive to success feedback and less responsive to failure. This results in a style of problem solving that causes the retarded individual to become a 'failure avoider' (Cromwell 1967). They are more motivated to avoid failure than to achieve success and their inferior performance reflects not so much a cognitive inadequacy as weak motivation (Gardner 1957; Zigler 1966 ; Zeaman and House 1963; Zigler and Harter 1969).

It can be predicted that lowered expectancy of success is also likely to influence the approach of mentally handicapped children to assessment situations.Indeed, motivational factors might affect the performance of this group of children to an even greater degree than has been implied in relation to non-handicapped children.

If it is accepted that there is a need to review current methods of assessing the cognitive abilities of handicapped children in the light of tendencies such as failure avoidance, there is an even more pressing need to investigate the potential effects of these same motivational factors on the children's cognitive development. The study of the *developmental* role of motivation - the so-called 'motor' of development in normal children - has been neglected in mental handicap research. Despite the recognition that motivational variables do affect the demonstration of competence in mentally handicapped adults (Cromwell 1967, Zeaman and House 1963) and children (Balla et al 1971, Balla 1980), there have been few attempts to study the developmental origins of these effects. It is not commonly recognised that the development of cognitive deficits in the handicapped child might result as much from motivational problems as from intrinsic functional anomalies.

It is easy to see how repeated failure as a result of intrinsic deficits in functioning could influence general expectancy of success (or failure) and thus adversely affect motivation to succeed. Seligman's (1975) 'learned helplessness' model encapsulates the essence of this problem. According to this model, a state of learned helplessness is reached when an individual perceives that s/he lacks control in obtaining a desired outcome. The major consequence of this is a lowering of expectations and a reluctance to believe in control when it is restored. Seligman's original model was based on research with animals. The reformulated learned helplessness model of Abramson et al (1978) is however equally applicable to the human learning process: the mentally retarded person is viewed as essentially defeated by chronic failure to cope effectively.

In mainstream developmental psychology the learned helplessness model has recently received increasing attention (e.g. Dweck and Repucchi 1973; Dweck and Bush 1976; Andrews and Debus 1978; Diener and Dweck 1980; Craske 1988, Chapman 1988). It is accepted that certain groups of children do suffer from motivational problems that may lead to cumulative skill deficits, and that through appropriate intervention methods, it is possible to alleviate these problems and thus help the child to reach his or her full potential. In the field of developmental testing, it is also becoming accepted that motivation and competence "cannot be teased apart" (Hrncir et al 1985) - that the measurement of 'competence' with a standardised measure inherently also includes a measurement of motivation (see also McCall 1981; Scarr 1981; Scarr-Salpatek 1986).

It is now widely recognised that there is a continuous interplay between cognitive and motivational factors in the development of competence in the non-handicapped child. If a child is seen to be underperforming, this is attributed to a maladaptive motivational pattern, or to a poor academic self concept. The major learning problem manifested by this child is a motivational one. Although the mentally handicapped child may by definition experience difficulties which can be attributed to functional anomalies there is no reason to assume that motivation plays only a secondary role in these difficulties. The learning process in mental handicap may be as much impeded by motivational deficits as it is by the cognitive deficit. Clearly, there is a need for an approach to the study of cognitive development in mental handicap which focusses as much on the relationship between performance and competence as on the sequential and structural similarities in the development of handicapped and normally developing children.

It could indeed be argued that the expanded developmental perspective now applied in mainstream psychology would benefit from the study of non-cognitive factors in development in children with mental handicap. It may be that that a full understanding of the interplay between cognition and motivation in development could best be achieved through the detailed study of a population in which an imbalance between the two may be expected. The high level of motivation characteristic of normal development is presumably sustained by experience of success and failure in satisfactory proportions. In contrast, by nature of his or her handicap, a child with mental handicap must have an increased experience of failure. This in turn might be expected to adversely affect the child's motivation to learn or even to demonstrate what has already been learned. Careful observation of handicapped children of different ages in both learning and assessment situations might therefore provide some insight into whether any deficit in motivation is manifested in identifiable patterns of behaviour, whether its effects on development are cumulative, and whether intervention might to any extent prevent this deficit from adversely affecting the developmental process.

While this thesis focusses on development in children with Down's Syndrome, it is hoped that findings may be of relevance to the field of mental handicap in general. It is important to state, however, that it cannot be assumed that all findings from research into DS can be generalised to mental handicap in all its forms. The fact that there are additional complications associated with the syndrome immediately sets it apart from other forms of mental handicap in which no such additional problems interfere with development. Nonetheless, at the very least, conclusions from studies of DS may be of value in directing the focus of research with other mentally handicapped - and indeed, non-handicapped - populations.

In an attempt to achieve a more integrated picture of the cognitive development of the DS child this thesis will present a detailed investigation of the relationship between competence and performance, with particular attention being paid to the non-cognitive aspects of performance in contexts of ostensibly cognitive tasks. The performance of DS and non-handicapped children on a number of differing cognitive tasks will be compared at a number of different ages with the aim of exploring the extent to which motivational factors may be impeding the development of cognition in the child with DS.

CHAPTER TWO

INCREASING TASK MOTIVATION IN DS CHILDREN: A COMPARISON OF TRIAL-AND-ERROR AND ERRORLESS TEACHING PROCEDURES

Motivation and learning

In the previous chapter it was suggested that a fuller picture of the cognitive development of the handicapped child might be achieved by broadening the scope of investigations to include the non-cognitive characteristics of performance in cognitive contexts, and by examining the way in which these may interfere with the development of cognitive ability. It was further suggested that the cognitive deficits seen in the DS child may to some extent result from motivational problems: that frequent early experience of failure may erode motivation to learn, thereby contributing directly to subsequent deficits in functioning.

The work of Dweck and her colleagues with non-handicapped children has underlined the importance of taking motivational factors into account in investigations of learning difficulties. Dweck and Repucci (1973) demonstrated that one group of children in their studies was more likely to give up in the face of failure than to persevere. They observed that these 'helpless' children tended to place less emphasis on the relationship between effort and success and when unsuccessful, to attribute the outcome of their behaviour more often to lack of ability rather than to lack of effort. The implication of this is that 'helpless' children see themselves as less instrumental in determining outcomes and are therefore less likely to respond to failure with increased effort or perseverance. These different attitudes were seen to emerge in spite of the fact that the 'helpless' and 'non-helpless' children were of similar levels of cognitive ability. Intervention techniques were therefore devised on the basis of altering the helpless children's perceptions of the reasons for their failure.

In the study to be presented here, the effectiveness of two strategies for teaching discrimination to DS and non-handicapped children were compared - trial-and-error and 'errorless' learning. As in the Dweck studies the aim of this experiment was to investigate a possible means of increasing the child's motivation to perform to full competence. Dweck's approach involved changing childrens' perceptions of the reasons for failure. This approach may not however be appropriate with mentally handicapped children. In comparison to non-handicapped children, the imbalance in the handicapped child's experience of failure does result from a lower level of ability. It is nevertheless still important that such a child does learn to recognise and to maximise the potential s/he does have. This is particularly crucial if an attitude of learned helplessness is to be prevented from reducing effort and perseverance in all learning situations, easy or difficult. It is equally important that handicapped children also do learn to attribute success to effort, instead of consistently opting out of challenging situations because they perceive their lack of ability to be an insurmountable obstacle to the achievement of success. Moreover it is crucial that handicapped children remain motivated to *sustain* the level of effort required for the achievement of this success.

The role of errors in learning

Many learning theorists - in particular Piaget (1936,1937,1950) - would argue that erring is an essential element in the learning process and that the experience of erring has as much to contribute to the process of cognitive development as does success. From observations of non-handicapped children in learning situations has come the acceptance that, in certain types of learning situation, the advent of error can often be a sign of progress. Mentally handicapped children, by comparison, have frequently been observed to avoid situations in which they might experience failure. It could be argued, therefore, that this very tendency may make a significant contribution to difficulties encountered in learning. If failures are not responded to, they cannot be incorporated into the accommodation/assimilation sequence which Piaget considers to be the trigger for progression onto higher stages of development. John Morss (1979), in a study of the development of the object concept in a group of DS children, observed that their learning did appear to be 'incomplete' in this sense. Initial successes seemed to be viewed in limited terms - more in terms of success on a particular occasion rather than as reflecting a general understanding. Morss offers this as an explanation for the inconsistency in performance demonstrated by many of his subjects; they did not seem to be responding to errors as a means of preparation for later successes.

It is not difficult to see how a negative attitude towards erring could evolve in the mentally handicapped child. Whereas a non-handicapped child is occasionally rewarded during the process of trial-and-error learning with sufficient levels of success to counteract the effects of inevitable errors, for the child with a mental handicap, the balance is more likely to be in the direction of consistent failure. Rather than recognising the positive value of errors and subsequently attempting to use them to advantage, it may be that these children see their errors as an underlining of their lack of ability to cope and therefore reject them. This may explain the 'incompleteness' of the learning process: understanding of success is prevented by a lack of understanding of failure and a 'failure set' may be generalised to areas of comparative strength

Errorless learning

The errorless learning technique arose out of the claim, frequently made by proponents of the cognitive approach to mental handicap, that central to the condition is a lessened ability to manipulate the environment and therefore to learn spontaneously from it. Zeaman and House (1963) have argued that the failure of so many mentally retarded people to learn can be viewed in terms of a stimulus control model. On the basis of this model, failure in discrimination learning, for example, is interpreted in terms of failure to attend to the relevant dimensions of a stimulus. A major implication of this is that to improve performance it is necessary to obtain greater control over the discriminative stimuli.

As noted above, the Piagetian view is that over the course of trial-and-error learning, as much is learned from mistakes as from successes. Terrace (1963a & b), working with pigeons, questioned this assumption and showed that learning is in fact possible without the subject ever having to respond to the negative stimulus: in short that learning *was* possible in the absence of errors. A stimulus fading procedure was used in which the alternative stimuli were gradually 'faded in' by increasing their prominence along a particular dimension (such as intensity, duration or size). In his original experiments, for example, Terrace taught pigeons to peck in the presence of a red key and to withhold pecking in the presence of a green key. On initial presentation the green key was less bright and of shorter duration than the red key. The brightness and duration of the green key were progressively increased over trials until they reached the same levels as that of the red key. Throughout all trials responses to the correct stimulus were reinforced and in the event that the incorrect key was chosen (the probability of this being far less than under trial-and-error learning), this response was neither positively nor negatively reinforced.

Later application of similar techniques with human subjects demonstrated that humans also are capable of learning in the absence of error (Moore and Goldiamond 1964). Success with non-handicapped children led to the growth of interest in this technique of 'fading' as a training method for teaching discriminations to mentally handicapped people (Sidman and Stoddard 1966; Touchette 1968; Sherman and Webster 1974). More recently, errorless techniques have been successfully used to teach very practical discriminations, skills previously unlearned when conventional methods were used (Cullen 1976; McIvor and McGinley 1983; Adams 1984).

Effective though it may be, the underlying assumption made in errorless learning is a very pessimistic one. Although to some degree its adoption implies recognition of the possibility that motivational factors may influence cognitive outcome, generally its use stems from the belief that the experience of erring in a learning situation will always be counter productive in the mentally handicapped, adding to already existing cognitive problems. Errorless learning is typically seen as a compensatory method for subjects who do not appear to be able to learn specific discriminations by trial-and-error.

On one hand, this approach may produce encouraging results from those who had previously been unable to make the target discrimination, particularly older children and adolescents. On the other hand, it is far from encouraging in terms of what it implies in relation to the ability to learn in mentally handicapped individuals i.e. that they are unable to learn from their mistakes. Evidence of poor learning ability using trial-and-error methods is not, however, in itself evidence that mentally handicapped children are unable to profit cognitively from their mistakes. There has been little direct evidence to support the view that such children cannot be taught to respond positively to error.

It follows from this that overuse of errorless learning techniques could potentially incur more harm than good. Encouraging learning through the use of success-only techniques could prevent learning taking place in situations where such a level of support is not available. If, as mentioned above, the advent of error can be a sign of progress, preventing the occurrence of error may prevent progress. Over-dependence on this sort of strategy could lead to a reluctance to advance beyond it. It may enhance the likelihood of correct response to one particular stimulus but may actually hinder development in other ways. Dweck, (1975) in describing errorless learning employed with a group of 'learned helpless' non-handicapped children, pointed out that;

> "Although there is ample evidence that errors per se do indeed have adverse effects on the performance of these children and that success works to motivate them, the question is whether the most effective way of organising such reactions to failure is to eliminate it from the situation, or to teach the child how to deal with it"
>
> *(Dweck 1975)*

If it is accepted that erring is an essential element in the learning process, it could be suggested that adoption of an exclusively, or primarily errorless approach may in some way pervert or undermine the normal course of early learning. It cannot nevertheless be denied that the errorless learning technique is an ingenious one that has met with a good degree of success. Careful use of such procedures may well be even more beneficial if a different emphasis were to be placed on their role in the learning process. In a number of experiments carried out with mentally handicapped subjects in the 1950s and 60s, for example, the differential effects on performance of prior experience of success and failure were examined. (Gardner 1957; Zeaman and House 1963). Results from these experiments generally indicated that although performance was likely to deteriorate following failure, after prior experience of success, performance was frequently enhanced. These results were achieved largely in experiments which investigated performance on a variety of motor tasks however. Although it was claimed that initial success did increase subjects' motivation to improve their performance, there appears to have been little scope for any generalisation to other, more cognitively based skills as a result of enhanced motivation.

The present study was designed to investigate whether an errorless technique could be used to increase motivation in a similar way, but with less restricted results. Rather than initially imposing a difficult trial-and-error learning task on an apprehensive learner, the intention was to use an errorless learning task to change the success/failure rate that would normally be first experienced on a learning task. In this way it would be emphasised to the child that s/he *could* learn and that learning could be easy, thereby raising motivation to put this newly realised skill to further use on a similar but more challenging, trial-and-error task. Two discrimination tasks were presented. The aim was to investigate whether success on the first discrimination task as a result of using the errorless teaching strategy might carry over into performance on the second, different task - which required trial-and-error learning.

By combining errorless and trial-and error-techniques, the aim was to set up a learning situation which would offer the handicapped child a more balanced experience of success and failure - one more similar to that experienced by non-handicapped children. Implicit in this is the assumption that such a strategy should have differential effects on the performance of handicapped and non-handicapped children. Devised on the basis of theory, errorless learning is an example of a teaching technique which, by emphasising the differences between mentally handicapped and normal children, may not have paid sufficient attention to any possible similarities between the two groups: development in handicap may not in essence be so radically different from normal development, but may appear so only because of an intervening variable (e.g. motivation). A direct comparison both of the effectiveness of errorless learning and of its potential role as a 'primer' in both DS and non-handicapped subjects is therefore required.

The two main hypotheses to be tested were therefore as follows:

1. That an errorless learning procedure would have a greater enhancing effect on performance than trial-and-error learning and that this effect would be greater in the DS group than in the group of non-handicapped children.

2. That initial experience of an errorless learning procedure would enhance performance on a subsequently presented trial-and-error task and that this effect would be greater in DS children.

STUDY 1

METHOD

Subjects: selection and matching

Both handicapped and non-handicapped subjects were selected and matched on the basis of inability to pass a subject selection pre-test based on the shape discrimination tasks which were to be used in the experiment itself (see below). To be included in the experiment it was necessary that children could demonstrate some knowledge of shapes but were as yet unable to discriminate the shapes used as target stimuli in the discrimination tasks.

In many studies comparing handicapped and non-handicapped children, subjects are matched on the basis of mental age. Given the particular nature and aims of the study to be reported here, however, it seemed important to avoid direct mental age (MA) matching for a number of reasons, both practical and theoretical. The validity of MA matching in handicap studies has frequently been questioned (Clarke and Clarke 1975; Woodward 1979; Wishart 1986), largely on the basis of its underlying assumption that subjects achieving similar test scores are equated on some fundamental intellectual dimension, regardless of the way in which those scores were achieved. The MA composite is arrived at by a simple addition of scores on a number of test items. Consequently it is possible for two testees to have identical MAs, but to have widely differing ability profiles. Even in studies of the normal population, some 'matches' must for some purposes be inappropriate.

Even more open to criticism are MA matches of mentally handicapped children to non-handicapped children of a much younger age; the greater the chronological age gap the more likely it is that the individual children's learning histories will have differed, both generally and in terms of specific skills. A single test item which tests only for the presence or absence of a particular skill can provide very little in the way of information about the developmental background to the acquisition of that skill. It cannot take into account the possibility that a child with a physical or mental disability may have taken considerably longer to have reached a specific developmental milestone than a younger non-handicapped child with an equivalent MA, or that in approaching that milestone their development may have taken different routes (see Chapter 1 p. 2). Nor can any account be taken of the possibility that throughout development the two children's experience of success and failure may have been very different, resulting in differing attitudes to learning in general and to the demonstration of what has been learned. Use of MA matching involves accepting that motivational factors do not differentially affect performance in the two populations and that performance and competence are similarly linked in handicapped and non-handicapped children.

Down's Syndrome group

Thirteen children were selected from two Edinburgh Special Schools. Three were excluded on the basis of the subject selection pre-test (see below) and a further two had to be dropped because of inability to concentrate for long enough to participate effectively in the experiment. This left eight DS children, five males and three females, who completed the experiment (mean age: 7 years 9 months; SD 14.6 months).

Non-handicapped group

Thirteen children were selected from two Edinburgh nursery schools. Four were already able to pass the subject selection pre-test and one had to be excluded later in the experiment because of attention difficulties. Mean age of the remaining eight non-handicapped subjects was 2 years 6 months (SD:4 months). There were four males and four females.

Design

The study was designed to investigate whether prior experience of learning using an errorless technique could be used to increase motivation to learn from a conventional trial-and-error teaching situation. Order of the two tasks was counterbalanced. Half of the children in each group (DS and non-handicapped children) were initially presented with the errorless learning task and then with the trial-and-error task, the other half with the two tasks in the opposite order (i.e trial-and-error followed by errorless learning) thereby enabling investigation of performance on a trial-and-error task without prior errorless 'priming'. An alternative design would have been to examine trial-and-error performance after exposure to another, similar trial-and-error task, and then to compare these scores with the trial-and-error scores of those subjects who had initially been presented with the errorless learning task. This design was avoided on the grounds that the intention of this study was to investigate a possible method for coping with failure - not to accentuate its effects or simply to confirm that DS children do not learn well in trial-and-error situations. The design adopted enabled observation of any differences in effects of the order of presentation of the two procedures in both the DS and the non-handicapped groups while also incorporating a control for any practice effects.

Procedure

Testing of all children took place in a small room, as free from distraction as could be arranged within the school settings. Child and experimenter sat opposite each other at a table.

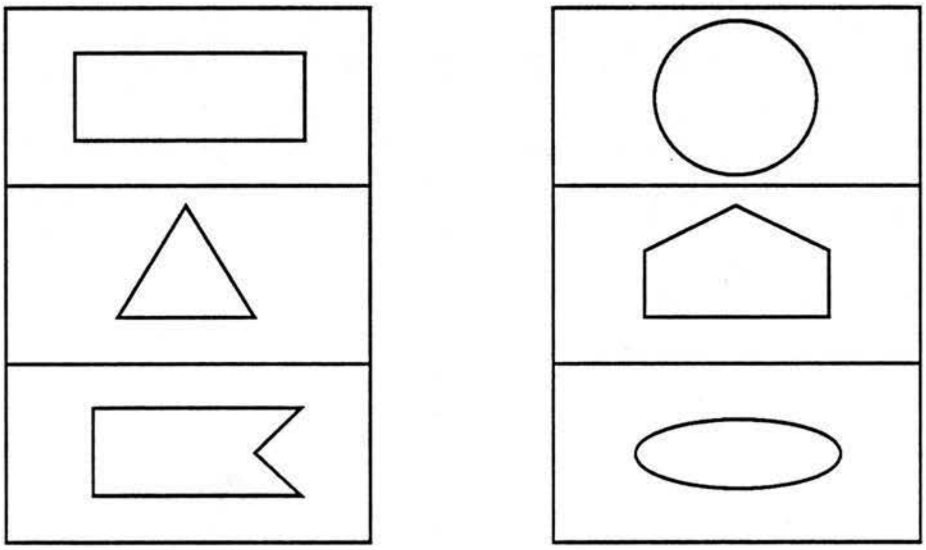
A total of 4 discrimination tasks were used, 2 shape discrimination tasks and 2 nonsense figure discrimination tasks. Nonsense figure tasks were included as a control for the possibility that prior experience with shape learning could have interfered with performance on the shape discrimination tasks. As it is impossible to control for any such differential experience of shapes prior to entering the study it was felt necessary to replicate the procedure using stimuli that no subjects could possibly have seen before.

All tasks were presented using a similar format. Each stimulus was presented on a separate card. In each trial three cards were placed, one by one, on the table in front of the child. Children were required to select the target stimulus from each of these trial sets of three.

Subject Selection Pre-test

In the selection pre-test one of two target stimuli (a rectangle or oval) was presented together with two alternative stimuli: one 'orthodox' shape (circle, square, triangle, parallelogram) and one of a variety of 'unorthodox' shapes introduced to limit the possibility that the oval or rectangle would be 'correctly' identified by a process of elimination of more familiar, known shapes (see Figure 2:1).

Figure 2:1
Test Cards Used in Subject Selection Pre-Test



All shapes were of different colours and approximately 7cm along the longest dimension. Each shape was centred on a white post-card sized card. Position of the shape to be identified was randomised over trials. Subjects were told "I am going to show you three cards with shapes on them. When I say the name of a shape, I want you to point to the card with that shape on it."

A mimimum of 28 trials were given to each subject, a minimum of six of each target stimulus and four of each orthodox shape. Where subjects' performance showed a bias neither towards correct or incorrect responses on the rectangle and oval trials, these were repeated until it could be established whether correct responses were due to random guessing or to a true ability to discriminate the shape in question. Responses were not differentially reinforced on this pre-test. As already stated subjects were selected on the basis that they showed an ability to discriminate some shapes but were as yet unable to discriminate rectangles and ovals.

The discrimination tasks

For both trial-and-error and errorless shape training the procedure consisted of 3 parts: a pre-test, the training trials and a post-test. No pre-test was required in the nonsense figure tasks as no subject could have had any prior experience of the target stimulus.

A/ Shape discrimination tasks

In the trial-and-error procedure an oval was used as the target stimulus; a rectangle was used in the errorless training strategy.

(i) Pre- and post-tests

Initially, 10 pre- and post-test trials were used for each shape. It was found, however, that this resulted in an unacceptably long procedure and the number of trials in each was therefore reduced to 7. The target stimulus (oval or rectangle) was varied in size and colour in these pre- and post-test trials in order to represent the concept of rectangle or oval in its more general form. The position of presentation in the row of 3 cards was randomised over trials. Instructions to subjects were as in the selection pre-test.

(ii) Trial-and-error training (oval)

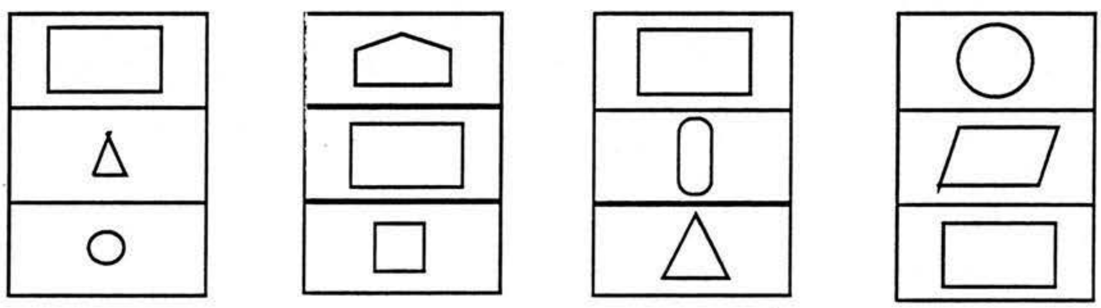
Trial-and-error training consisted of 15 trials. An oval was presented in every trial together with a common and an unorthodox shape of the same colour. Colour was changed over the 15 trial sets to prevent subjects from using this dimension as a discriminatory target; size and position of the ovals as well as order of presentation of the trial sets were randomised.

Subjects were told 'I am going to show you three cards with shapes on them just as I did the last time and I want you to point to the card with the shape I ask for. This time I will tell you if you are right." Correct responses were verbally praised. Incorrect responses were negatively reinforced; children were told: "No, that is not right. Try again the next time"

(iii) Errorless training (rectangle)

15 trial sets of three cards were used. The trials were presented in an order such that two alternative stimuli to the rectangle were gradually 'faded in', increasing in size while varying in shape over trials (see Figure 2:2).

Figure 2:2



In trial sets 1 and 2 the target stimulus was presented with two blank cards. In these and all subsequent trials, position of the target stimulus was randomised over trials, as was colour, the latter a precaution against misidentification of this as a relevant attribute of the target stimulus. Trial sets 3 and 4 consisted of a rectangle with two similarly coloured but much smaller (0.5 cm) alternative shapes. The dimensions of the alternative stimuli were increased over each set of two trials until trials 11, 12 and 13 where size was increased and colour changed over each single trial. These trials became increasingly more difficult as the sizes of the alternative stimuli became very similar to that of the target stimuli. The final two trials consisted of three red shapes each: a rectangle and two alternatives of equivalent dimensions.

Verbal instructions were identical to those given in trial-and-error training. Verbal praise was given each time the child made a correct response. Errors were not commented on but, rather than proceeding with the next trial, the previous trial was re-presented, this procedure being repeated as necessary until the child had shown mastery of that particular step in the training sequence.

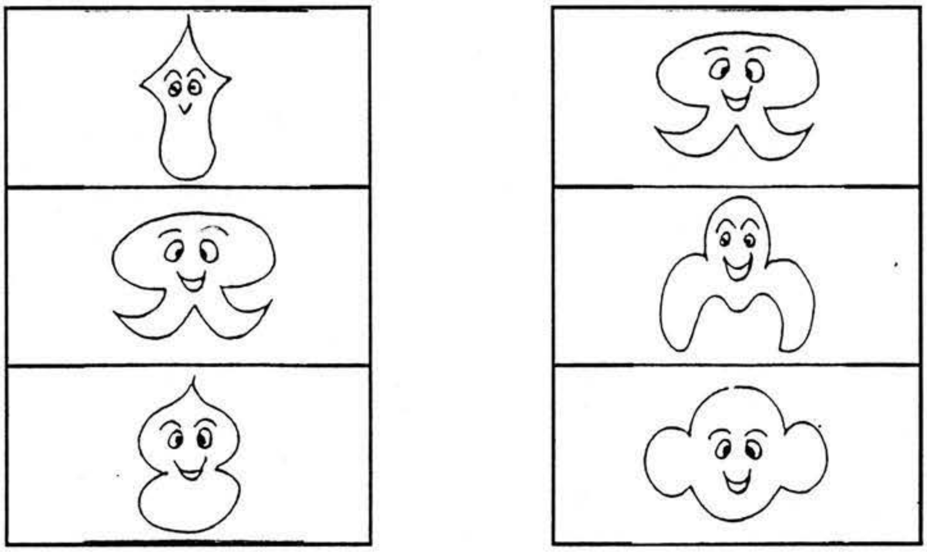
B/ Nonsense figure discrimination tasks

As in the shape discrimination task stimuli were presented on white post card sized cards. In trial-and-error training 'nims' were used as target stimuli; 'wugs' were used in errorless training (see Figure 2:3),

(i) Trial-and-error training (nims)

Subjects were shown a drawing of Mr Plimp and told that he had some friends called 'nims' in the pack of cards on the table. The aim of this game was to help Mr Plimp find all the 'nims'. Subjects were then told "I am going to show you three cards with little men on them. I want you to point to the card which you think has Mr Nim on it. Correct responses were verbally praised on behalf of Mr Plimp. When an incorrect response was made subjects were told to try again next time.

Figure 2:3
Examples of Trial and Error Learning Test Cards
(Nonsense Figure Discrimination — Nims)



(ii) Errorless training (wugs)

15 trial sets of three cards were used. Alternative stimuli to the 'wug' were faded in over trials in exactly the same way as in the errorless rectangle training.

(iii) Post tests

Post-test trials consisted of seven test card sets, each with either a 'nim' or a 'wug' and two other similarly sized nonsense figures. Subject instructions were as in the selection pre-test.

RESULTS

All responses in the pre-tests, training and post-tests were used in the analysis. Scores in the pre- and post-tests were expressed in terms of the number of correct responses made. Since the number of trials presented in the errorless procedure was subject-determined, scores in the errorless and trial-and-error training periods were expressed as percentages (correct responses/total responses).[1]

A/ Shape discrimination

Pre-test trials.

Pre-test scores on both shape tasks were compared in the DS and non-handicapped groups. No significant differences were found, thereby validating the procedure adopted for sample selection and matching (oval (trial-and-error): $t = 1.078$, df(14), NS; rectangle (errorless); $t = 1.56$, df (14), NS). There were also no differences within groups on the two discrimination tasks (DS: $t = 0.74$, df (7), NS; non-handicapped: $t = 0.42$, df (7), NS)

Training trials.

Table 2:1 shows the percentage of correct responses during trial-and-error and errorless shape training trials for both the DS and non-handicapped groups. In the combined groups performance in the two training conditions differed significantly ($t = 3.7$, df (15), $p < 0.005$). The same comparison in the two groups taken separately yielded a similar pattern in both groups (DS $t = 11.4$, df (7), $p < 0.0005$); non handicapped $t = 4.8$, df (7), $p < 0.005$). As was predicted, errorless training scores exceeded trial-and-error training scores in all cases, the effect being slightly more pronounced for DS than for non-handicapped subjects.

[1]These results have already been reported in the British Journal of Educational Psychology; L. Duffy and J.G. Wishart (1987). A comparison of two procedures for teaching discrimination skills to Down's syndrome and non-handicapped children, 57, 265-278.

T-tests comparing the trial-and-error training scores of children initially trained with errorless learning and those with no prior training revealed a significant difference in favour of those who had had prior errorless training ($t = 2.15$, df (14), $p < 0.025$). Again the same effect was present in the scores of the two groups taken separately, although in both cases it was slightly weakened. (DS: $t = 1.47$, df (6), $p < 0.10$; non-handicapped: $t = 1.63$, df (6), $p < 0.10$)

Table 2:1

Percentage of Correct Responses During Shape Training Trials

	Errorless Training (rectangle)	Trial-and-error Training (oval)
Initial Training	*DS subjects*	
Errorless	1) 100	60
	2) 100	46
	3) 100	66
	4) 100	33
Trial-and-error	5) 80	40
	6) 85	46
	7) 100	33
	8) 100	40
	Non-handicapped subjects	
Errorless	1) 82	60
	2) 94	60
	3) 89	46
	4) 100	100
Trial-and-error	5) 100	66
	6) 94	33
	7) 100	26
	8) 94	46

No overall effect of order of presentation of training conditions on errorless scores was found. In the DS group, however, an unexpected trend emerged: children who had initially been presented with the trial-and-error training task attained poorer errorless training scores than children who had had no prior trial-and-error training experience ($t = 1.70$, df (6), $p < 0.10$).

Table 2:2

Shape Discrimination - Effects of Errorless and Trial-and-Error Training on subsequent Performance: Pre/Post-Test Score Difference

	Errorless Training Scores (rectangle) Pre-	Post-	Diff.	Trial-and-Error Training Scores (oval) Pre-	Post-	Diff
Initial Training			*DS Subjects*			
Errorless	1) 0	5	5	2	5	3
	2) 0	7	7	5	7	2
	3) 4	7	3	2	7	5
	4) 2	4	2	1	6	5
Trial-and-Error	5) 1	2	1	1	2	1
	6) 0	5	5	2	5	3
	7) 1	5	4	2	1	-1
	8) 3	6	3	1	4	3
			Non-handicapped subjects			
Errorless	1) 1	5	4	4	7	3
	2) 3	6	3	2	6	4
	3) 2	6	4	2	3	1
	4) 6	7	1	7	7	0
Trial-and-error	5) 4	7	3	2	7	5
	6) 2	6	4	1	5	4
	7) 2	7	5	2	2	0
	8) 1	6	5	3	6	3

Post-test trials.

Improvement in performance was calculated by comparing pre- and post-test scores. Table 2:2 shows the differential effects of the two training conditions on post-test improvement. Overall a significant difference was found in favour of errorless training with performance improving more following training with this procedure than with trial-and-error training ($t = 1.78$, df (15), $p < 0.05$). This difference did not reach significance in either group however, although unexpectedly, it came closest to doing so in the non-handicapped group ($t = 1.76$, df (7), $p < 0.10$).

No overall effect of order of presentation of training conditions on post-test improvement was found. Only the trial-and-error scores in the DS group varied significantly with order of presentation, children with prior errorless experience producing better trial-and-error scores than those with no prior training ($t= 1.96$, df (6), $p < 0.05$).

In the two non-handicapped sub-groups performance was roughly equivalent in both training conditions, irrespective of their order of presentation.

B/ Nonsense figure discrimination

Training trials.

Table 2:3 shows errorless and trial-and-error scores. As in the shape task, errorless training scores reliably exceeded trial-and-error training scores. Again this difference in scores was less pronounced in the non-handicapped than in the handicapped group (overall $t = 6.15$, df (15), $p < 0.0005$; DS $t = 5.96$, df (7), $p < 0.0005$; non-handicapped $t = 3.31$, df (7), $p < 0.01$).

Comparison of trial-and-error training scores of children initially given the errorless task and children with no prior training showed no significant differences, either for the combined groups or for the groups taken separately. However there did appear to be a trend in favour of DS children with prior errorless training. This trend was not present in the trial-and-error training scores of non-handicapped children assigned to the same condition.

Table 2:3

Percentage of Correct Responses During Nonsense Figure Training Trials

	Errorless Training (Wug)	Trial-and-Error Training (Nim)
Initial training	*DS subjects*	
Trial-and-error	1) 94	40
	2) 100	26
	3) 82	13
	4) 94	80
Errorless	5) 100	33
	6) 94	46
	7) 100	40
	8) 100	86
	Non-handicapped subjects	
Trial-and-error	1) 100	33
	2) 100	100
	3) 94	66
	4) 100	73
Errorless	5) 100	33
	6) 94	80
	7) 94	53
	8) 100	73

There was no overall effect of training order on errorless training scores although, as in the shape task, there was a trend in favour of the DS sub-group given this training condition first i.e. with no prior trial-and-error experience, ($t = 1.48$, df (6), $p < 0.10$).

Post-test trials.

Since there were no pre-test scores, the differential effect of the two training strategies on nonsense figure discrimination was evaluated by comparison of post-test scores in the two training groups (see Table 2:4).

Table 2:4

Difference Between Nonsense Figure Post-Test Scores after Errorless and Trial-and Error Training

	Errorless training	Trial-and Error training	Diff.
Initial training		*DS subjects*	
Trial-and-error	1) 5	3	2
	2) 7	0	7
	3) 7	5	2
	4) 6	3	3
Errorless	5) 7	0	7
	6) 6	0	6
	7) 7	7	0
	8) 7	7	0
		Non-handicapped Subjects	
Trial-and-error	1) 7	7	0
	2) 7	7	0
	3) 6	3	3
	4) 7	7	0
Errorless	5) 6	3	3
	6) 6	5	1
	7) 6	4	2
	8) 7	7	0

Errorless learning proved superior in all comparisons with the effect again more pronounced in the DS group than in the non-handicapped group (overall results: $t = 3.6$, df (15), $p < 0.005$; DS group: $t = 3.2$, df (7), $p < 0.01$; non-handicapped group: $t = 2.18$, df (7), $p < 0.05$).

Although a trend in favour of initial errorless learning was evident in post-test scores of both non-handicapped and DS groups, these differences failed to reach significance - either overall or in either subject group.

DISCUSSION

This study aimed to investigate and compare the efficiency of errorless and trial-and-error methods in teaching discrimination skills to non-handicapped and DS children. As predicted, in both shape and nonsense figure tasks, errorless learning training scores reliably exceeded those produced under trial-and-error training. This overall pattern is consistent with previous findings from studies using errorless learning techniques. As a strategy for teaching specific within-task discriminations, the technique is clearly more effective than trial-and-error methods. In addition, the prediction that this difference would be more pronounced in the DS children's scores was supported. This is hardly surprising. The procedure in the trial-and-error task after all most closely approximates the conditions encountered in everyday natural learning situations and the handicapped children by definition show inferior learning skills in such situations. In comparison, performance of non-handicapped children would not be expected to differ to such an extent in the two types of training trials.

More surprising, however was that the relatively superior response of DS children to errorless training in the shape task did not carry over into post-test improvement. Unexpectedly, the errorless procedure seems to have benefited the non-handicapped group's post-test performance more than that of the handicapped group. This may be explained by an interesting pattern which emerged from an investigation of the effects of order of presentation of the two tasks. In trial-and-error training, all subjects (both DS and non-handicapped) who had received prior errorless training produced better scores than those who had had no initial training. By contrast, the only *difference* in errorless training results was found between the two DS sub-groups, with those subjects who had previously been presented with the trial-and-error task producing worse scores; non-handicapped subject groups showed no such order effect. This same pattern was repeated, albeit to a lesser extent, in the nonsense figure task.

Comparison of the mean errorless shape training scores achieved by non-handicapped subjects with and without prior trial-and-error training in fact indicates a trend in favour of those who *had* previously been presented with the trial-and-error task. Overall, scores achieved by the non-handicapped group on both tasks - errorless and trial-and-error - were higher where presentation of the task had followed another task - possibly indicating the operation of a practice effect.

The reverse of a practice effect seems to have taken place in DS errorless training performance on the shape task. Prior trial-and-error experience seems to have had an adverse effect on their performance on the errorless task. Poorer training performance may therefore have been responsible for reducing post-test scores and the consequent narrowing of the overall difference between the amount of improvement made under the two training conditions. An interpretation of this pattern could be that initial errorless training produced the same enhancing effect on trial-and-error training scores in the two groups but for different reasons. Non-handicapped children had simply become better practised in discrimination tasks, whereas the success-only support given by the errorless training had led to improved DS performance in the subsequent trial-and-error task.

These findings lend support to the hypothesis that motivational factors differentially affect the expression of performance and competence in the two groups. It is significant that for the DS group, even in a structured learning situation such as that provided by the errorless strategy, prior trial-and-error training *does* seem to have had a negative effect on performance. Errors inevitably encountered in the trial-and-error task may have lowered expectations of success in the subsequently presented task, thereby reducing motivation to perform to full potential on that task.

No such apparently adverse effects of prior trial-and-error training were evident in the post-test performance of the non-handicapped group. This is not especially surprising; given the superior ability and better balanced history of success and failure in this group, there is no reason to expect that their motivation to learn would be reduced by trial-and-error experience. Nor indeed would it be predictably increased to any significant extent by errorless experience.

In the shape task the outcome of this overall pattern was that it appeared to have masked the effect that prior errorless training *did* have on the subsequent trial-and-error performance of children in the DS group. Whereas there was no overall carry-over effect, post-test improvement was markedly more evident in the DS sub-group who had not had previous experience of the trial-and-error task. If this result can be attributed to increased motivation brought about by success-only experience in the errorless task, it is encouraging, particularly given that only a single training session was required to produce such an effect.

In contrast, the effect of errorless training on nonsense figure post-test scores was immediately apparent. Although this effect was more pronounced in the DS group, both handicapped and non-handicapped children in fact appeared to have benefited from errorless training to a greater extent than from trial-and-error training. The fact that differences between the two groups were not as marked in this second set of tasks may highlight the effects of prior learning experience on performance on the shape tasks. Neither group could have had any prior learning experience of the nonsense figures; however it had been a prerequisite for participation in the shape study that some knowledge of shapes was already present. The overall age difference between the two groups, in conjunction with their similar pre-test abilities would appear to indicate that their learning histories prior to training must have differed, the DS group having been exposed to a higher absolute and relative rate of failure than the non-handicapped children. Poorer performance on the shape discrimination task may therefore be associated with the fact that DS children were being tested on a concept in which prior experience of failure was actually being added to, further lowering expectations of success - even on the easier errorless task.

Overall, this set of results contains a consistent trend. Prior errorless training was found to have enhanced the subsequent trial-and-error performance of children in the DS group on both sets of tasks. Although this effect was only statistically significant on the shape task, nevertheless it did systematically affect performance in the same positive direction in both tasks. Moreover there was very little evidence of any adverse effects of prior trial-and-error training on performance in the non-handicapped group. The presence of such effects in the DS children's scores however, underlines the importance of ensuring that expectations in the handicapped do not lie in the direction of failure.

Although in all cases errorless performance was superior to that achieved under trial-and-error training, it must be borne in mind that this was not consistently reflected in a subsequent test of what had actually been 'learned' in training, i.e in post-test results. Errorless procedures have been previously criticised on the grounds that although effective at a within-task level (Schimoeller et al 1979; Etzel et al 1981; Gollin and Savoy 1968), there is little generalisation to other tasks. As a teaching technique per se, it could be maintained that errorless learning may have little intrinsic value, its usefulness limited to situations in which very specific responses are required which have failed to be learned by trial-and-error methods. However, the results from this study indicate that by virtue of the priming effect they appear to have on trial-and-error performance, errorless techniques may be useful as a means of increasing the motivation of children to make a greater effort in everyday, trial-and-error learning situations. Dweck (1975), as noted previously, has stressed that children should be taught to recognise the importance of effort, not just ability, as a determinant of success and failure. In contrast to the findings from this study she found no positive effects from an errorless - or success-only - training strategy on the performance of a group of 'helpless' children on a subsequent test. She found that a more effective method of altering these children's perceptions of success and failure was to teach them more directly to attribute success to effort. This contrast may result from the nature of the two subject groups. Both the DS children in the present experiment and the 'helpless' children in Dweck's studies seemed to be affected by a similar reluctance to perform to full potential, but the actual levels of potential in the two groups must inevitably have been different. A non-handicapped child who is taught to realise the more favourable outcome of increased effort will come to such a realisation through frequent reward for that effort i.e. through increased success. The handicapped child - by definition - is unlikely to experience a sufficiently high ratio of success and failure to justify such perseverance.

Errorless learning may therefore be a more appropriate means of dealing with learned helplessness in mentally handicapped individuals.

Rather than simply avoiding potential learning situations as non-handicapped 'helpless' children appear to do, they are at the added disadvantage of being more likely to fail in situations where they do choose to apply themselves - thereby reinforcing their lack of belief in their ability to succeed. As a consequence of this they may need constant reminding that success is possible.

Some anecdotal evidence from the present study can be offered in support of the contention that it is important to avoid the establishment of a 'failure set' being generalised to areas of comparative strength. Patterns of performance in at least two of the DS children - both 10 year olds - seemed to suggest that they had underperformed in the selection pre-test. One of these children in fact readily admitted afterwards that this had been the case. Two other DS 10 year olds also had to be excluded from the analysis because although they produced overly consistent incorrect responses in the selection pre-test, when offered tangible reinforcement for correct responses they changed to near-perfect levels of success. According to the school teacher, in one case at least, this behaviour was likely to have been a deliberate tactic, produced to avert the possibility of being presented with a more difficult task.

This reluctance to perform to full potential was interestingly only evident in the older DS subjects. A 6 year old boy with DS who had to be excluded from the study on the basis of high performance on the selection pre-test made no attempt to hide his capabilities in this test. By comparison, the older children, rather than allowing themselves to be placed in a situation over which they might have no control, seemed to be imposing their own control over the situation from the start, their poor performance very much a case of "won't do", rather than "can't do" (Koegel and Mentis, 1985).

Although the numbers of children on which these observations were made were small, the differences in attitude demonstrated by older and younger DS subjects do resemble those which would be expected within the learned helplessness model, with the effects of failure becoming more prominent as age - and therefore experience of failure - increases. Given this it seems important to attempt to determine whether this progression can in any way be inhibited. As suggested previously, it may be that the functional anomalies intrinsic to the condition of handicap are only partially responsible for the extent of the accompanying cognitive deficit. Persistent failure in itself may to some degree be responsible for the development of a negative attitude toward learning, compounding the already-existing disadvantage of the mentally handicapped child in learning situations.

If it were possible to identify the developmental origins of this tendency to avoid potential failure situations, early intervention programmes could be targeted towards preventing its emergence, or at least limiting its development. It cannot simply be assumed, however, that the motivational deficit does evolve simply as a result of repeated experience of failure. It would, nevertheless seem important to investigate this possibility. Motivational problems may actually be intrinsic to the condition of handicap, but even if this is the case, this should not prevent attempts to minimise their effects. Through careful observation of early cognitive development, it may be possible to determine at what stage motivational deficits are first manifested and to identify the most appropriate means of trying to counteract their effects. The study to be presented in the following chapter therefore focuses on the very earliest stages in DS development, infancy.

CHAPTER 3

A CROSS-SECTIONAL STUDY OF THE PERFORMANCE OF CHILDREN WITH DOWN'S SYNDROME ON THE BAYLEY SCALES OF INFANT DEVELOPMENT

Findings from the experimental study reported in Chapter 2 indicated that the performance of DS children on a discrimination task was affected by non-cognitive factors to a greater extent than was that of non-handicapped children. Not only did prior error-free experience have an enhancing effect on DS performance on a subsequently presented trial and error task, but initial trial and error training adversely affected subsequent errorless performance. Both patterns would appear to indicate that the DS childrens' approach to discrimination learning was to some degree dependent on the levels of success and failure experienced in the learning situation. Neither pattern was present to a similar extent in the scores of the non-handicapped group, nor was there any evidence that any of the non-handicapped children were underperforming in any of the trial sets.

While suggestive, findings from this one study can obviously tell us very little about whether - and to what extent - non-cognitive factors may have affected learning at earlier stages in the development of these DS children. They can also tell us little about how motivational factors might have influenced acquisition of other types of cognitive skill. Clarification of such issues would require examination of cognitive behaviour at much earlier stages in DS development. In the study to be presented in this chapter therefore, a group of 36 children with DS aged from 3 months to 5 years were presented with a wide-ranging set of cognitive tasks and both the qualitative and quantitative aspects of their performance investigated. The assessment battery used was the mental scale of the Bayley Scales of Infant Development (BSID). The items contained in psychometric tests of infant development such as the BSID provide a 'checklist' of the sorts of skills that are commonly seen to emerge as cognitive ability develops.

The BSID is designed to assess:

> sensory-perceptual acuities, discriminations and the ability to respond to these; the early acquisiton of 'object constancy' and memory, learning and problem- solving ability; vocalisations and the beginnings of verbal commmunication; and early evidence of the ability to form generalisations and classifications, which is the basis for abstract thinking.
>
> *(Bayley 1969)*

It provides a single outcome measure which can be used to determine the ability of a child relative to that of other children and of groups of children with handicaps of varying aetiologies relative to the normal population. In the study to be presented performance on the Bayley will be examined at two levels. The test scores collected will provide a quantitative cross-sectional picture of the general cognitive ability of this group of children with DS. Behaviour during testing will also be monitored, however, for any more qualitative evidence of levels of performance being influenced by non-cognitive variables; parents' reports will also be collected for this purpose. The primary aim of this study will be to determine whether it is possible to identify the developmental origins of the motivational deficit observed among older DS children in the experimental study in Chapter 2.

STUDY 2

METHOD

Subjects

Thirty-five subjects were drawn from a volunteer subject pool run by the Department of Psychology and recruited through the local health authority and the Scottish Down's Syndrome Association. This subject pool comprises approximately 90% of all children born with DS in the Lothian area and previous research has shown that it is a representative sample in terms of variables such as social class, level of parental education, ability level and secondary health problems. All children aged between 3 months and 5 years at the time of the study were included. Table 3:1 shows the number and sex of children tested in each age group.

Age (months)	N	Male	Female
3	8	5	3
6	5	1	4
9	2	0	2
15	1	0	1
18	1	1	0
21	1	0	1
24	2	1	1
30	2	0	2
36	4	3	1
42	4	1	3
48	2	1	1
54	2	2	0
60	1	0	1
Total	**35**	**15**	**20**

All subjects were confirmed by karyotyping as having Down's Syndrome (DS); all were standard Trisomy 21. Fourteen had congenital heart disease of varying degrees of severity (the majority minor), 4 had diagnosed hearing loss, 10 had minor visual impairments and one suffered from arthritis. Fourteen were first borns. Numbers of siblings at time of testing ranged from none to 4.

Each child in the present study served as his or her own control. The BSID has already been standardised on a large group of non-handicapped children and since the purpose of this investigation was to observe test performance in DS children, no direct control group was considered necessary. It will be recalled that the practice of mental age matching of handicapped and non-handicapped groups was avoided in the study presented in Chapter 2 for practical and theoretical reasons. All of these reasons are equally applicable to the present study. It was in fact hoped that findings from this study would provide additional empirical evidence to support the viewpoint that mental age matching is of questionable value in studying development in handicapped populations.

PROCEDURE

All testing took place in the Department of Psychology in a small quiet room with a minimum of distractions. Testing was scheduled for a time of day when the mother thought her child was most likely to be alert and cooperative. The mother was present throughout and the child sat on her knee for testing; mothers were asked to encourage their children if necessary but not to direct their responses in any way. They were however asked to comment on their child's performance - in particular when test items were failed - and to indicate whether s/he had previously been observed to produce the behaviour required by a given test item in any other setting.

The BSID was administered in accordance with the procedural instructions provided in the test manual (Bayley 1969). The 163 items in the test are arranged in order by age placement (the age at which 50% of the standardisation sample passed a given item). Each item has also been assigned an age range value which is an estimate of the ages at which they were passed by 5% and 95% respectively of the children tested (for example, item 88 - picks up cup: secures cube - is placed at 9 months, but has an age range of 6 - 14 months). Generally, a child's test performance is expected to extend over items having age placements several months apart. Basal and ceiling levels determine the upper and lower limits of performance, the basal level being the item preceding the earliest failure and the ceiling level the item representing the most difficult item on which the child succeeded. To establish basal and ceiling levels a criterion of 6-10 successive passes or fails is recommended.

Although cognitive development in children with DS is generally described as progressing at half the normal rate (see e.g. Berry et. al 1984), it was necessary to establish more accurately for each individual subject the most appropriate level at which to start. Basal levels were determined on the basis of observations of the child playing for a short period before testing and by using the Situation Codes provided in the manual. These Situation Codes refer to series of items in which one basic test procedure or stimulus situation can be used for several items of varying difficulty levels. One particular task, for example, which involves placing square and circular blocks onto a form board (referred to in the test as the blue board) covers 7 different items, with age placements ranging from 13.6 to 30+ months. A child's performance on this task provides a useful estimate of the age level at which to begin testing.

DATA ANALYSIS

BSID scores were analysed in a number of ways, outlined below. In addition records of the more qualitative aspects of children's behaviour during testing were examined for any indication that motivational factors were in any way influencing performance. Parents' reports were also used for these analyses.

Treatment of scores

Bayley scores can be expressed in several different ways. The conventional measure is the Mental Development Index (MDI). This is converted from the raw score (the total number of items passed) and is expressed as a standard score with a mean of 100 and a standard deviation of 16. A Developmental Age Equivalent or Mental Age (MA) in months can also be calculated from the norms tables by comparing raw scores with corresponding MDIs of 100 at different age levels. A fourth measure is the developmental quotient (DQ) based on the ratio of developmental to chronological age.

There are a number of advantages and disadvantages associated with each of the above measures. Because MDIs are based on a normal distribution of scores they are easily interpreted as a measure of development. The normative tables which cover the age range from birth to 30 months contain Developmental Indices ranging from 50 to 150 (approximately 3 sds below and above the standardisation mean). Use of MDIs are therefore precluded when raw scores fall below an MDI of 50 at any age level. Extrapolated values for raw scores falling below this level have been computed (Naglieri 1981) but because these were arrived at by calculating regression equations from the normative sample, they can only be used with children of 30 months or below (the age range of the BSID standardisation sample). For the purposes of the present study, which includes children of up to 60 months of age, these scores are therefore not appropriate. Naglieri's values are not in any case based on actual empirical data with the result that some precision is inevitably lost.

The most precise measure of performance can be achieved by expressing scores simply as the total number of items passed i.e as raw scores. These also permit direct comparison of performance in infants whose scores are too low to be converted to MDIs using standardisation sample norms. Raw scores are rarely used in the literature, however, probably because, unlike MDIs, they are not so readily interpreted as a measure of development. Mental Age scores, like MDIs also have the advantage of allowing evaluation of development in readily understood terms. It is not possible to obtain an MA for every raw score, however, which makes this the least precise of the four methods of measurement. DQs present similar difficulties when used alone. Despite the fact that Bayley herself stressed that there is no evidence to support interpretation of figures of this kind derived from the BSID, this is one of the most frequently used means of expressing Bayley scores. Use of DQs permits the aggregation of data from different age groups but by their inherent nature, such scores can increasingly penalise handicapped subjects as they grow older, suggesting deceleration in development even when developmental rate is constant (Bailey and Bricker 1985).

In the present study each of the four different means of expressing scores (MDIs, raw scores, MAs and DQs) will be used in analyses depending on which is most appropriate for the level and type of analysis being carried out.

Analyses

Table 3:2 and Figure 3:1 present the mean raw scores achieved at each age level together with the mean number of items passed at these ages by the BSID standardisation sample. It can be seen that although in general, DS raw scores do increase between every adjacent pair of ages, the difference in mean raw scores between the two populations widens as age increases. This is further demonstrated in the rising number of DS scores which fall below MDI levels of 50 as age increases. At 3 months, all 8 DS subjects attained MDIs of 70 or more, with 4 scoring above 80. One 6 month score fell short of an MDI level and had to be converted into a mental age score, but the MDIs of the 4 remaining 6 month olds ranged between 60 and 81. Both 9 month old subjects did achieve sufficiently high raw scores to permit conversion into MDIs but these were lower than almost all MDIs attained by the two younger age groups. The highest MDI of 98 was attained by the only one 15 month old in the sample; her raw score exceeded that of the single 18 month and 21 month old subjects, both of whom were able only to produce mental age scores rather than MDIs as was one of the two 30 month olds. The two 24 month olds' MDIs both stood at 54. As norms are only provided for children up to 30 months, all scores beyond this age had to be converted into Mental Ages.

Table 3:2: Mean raw scores

Age (months)	N.subjects	Mean raw score for age level	BSID norm
3	8	26.9	36.7
6	5	52.4	71.8
9	2	80	10m 83.0
15	1	113	113.8
18	1	101	125.4
21	1	104	137.0
24	2	119	144.6
30	2	125.5	154.4
36	4	131.75	-
42	4	137.75	-
48	2	151.25	-
54	2	155	-
60	1	161	-

Figure 3.1: Mean Raw Scores

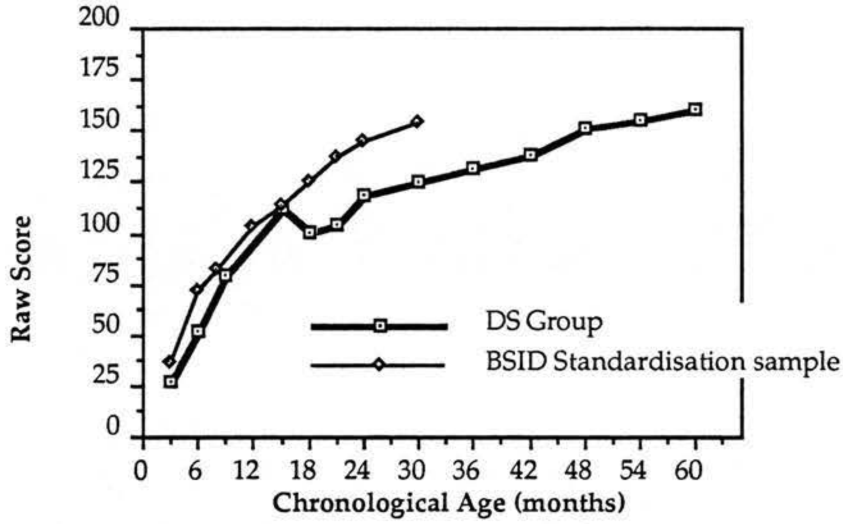


Table 3:3 shows the MA scores for all subjects. This table, like Figure 3:2, demonstrates that although mental age does continue to increase, it proceeds at a much slower pace than that seen in normal development where chronological and mental ages are roughly equivalent. The widening gap between the two populations in terms of rate of development is perhaps more clearly expressed through direct comparison of chronological and mental ages: MA lags behind CA by only 0.5 months at CA 3 months, increasing to a difference of 12 MA months at 12 months of age and to 24 months at CA 54 months.

Table 3:3: Mental Ages

Age (months)	Mental Age (months)	Difference between CA and MA
3	2 - 2.5	0.5 - 1
6	4.5	1.5
9	7 - 8	1 - 2
15	14 - 15	0 - 1
18	11 - 12	7 - 8
21	12 - 13	11 - 12
24	16 - 17	7 - 8
30	18	12
36	20	16
42	21 - 22	20 - 21
48	26	22
54	30	24
60	30+	30

Figure 3.2: Mean Mental Age Score

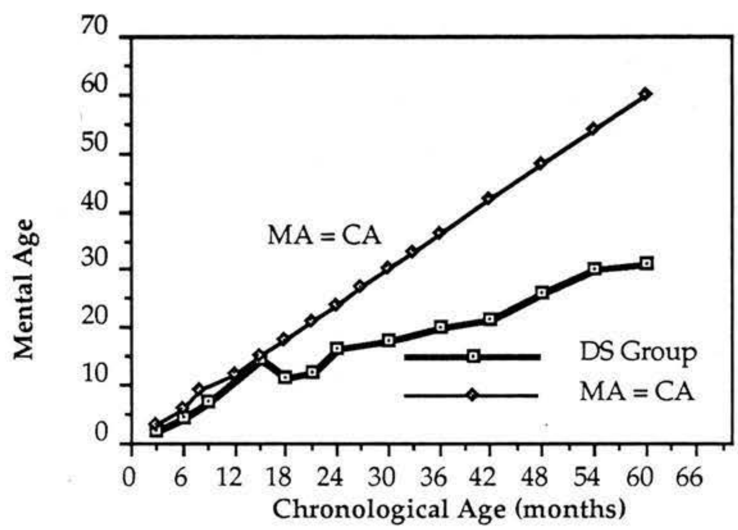
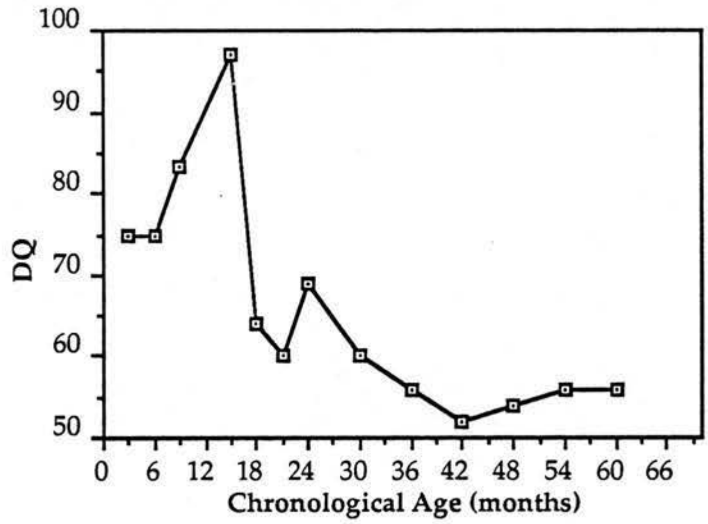


Figure 3:3 presents the decline in the slope that results when MA/CA measures are translated into Developmental Quotients. The fact

that the path of the curve is slightly interrupted by the high score achieved by Subject 15 (15 months) and the comparatively lower level of performance attained by Subject 17 (21 months) is indicative of the tremendous variability existing between individual subjects both within and across age groups in levels of functioning.

Figure 3.3: Mean DQ Scores



Because of the small numbers of subjects in several of the mid-range age groups, it is not possible to identify any clear pattern of increasing or decreasing variability in scores with increasing age. Mental ages attained by 3 month old subjects range between below 2 months and 2.5 to 3 months; at 6 months a difference of 2.5 MA months was found between highest and lowest scoring subjects; the two 30 month old subjects achieved scores 4 to 5 MA months apart; at 36 months this had increased to a difference of 6 to 7 MA months; at 42 months MAs ranged between 18 and 30 months. Overall these figures would seem to indicate that inter-individual variability of performance rises with increasing age. Given that as MA increases, there is more scope for individual differences, this pattern may be artefactually exaggerated however.

An apparent deviation from this pattern was found at 24 months where the two subjects in this age group (Subjects 18 & 19) attained identical raw scores. Closer analysis of their score profiles, however, revealed that these scores were arrived at in very different ways. The most striking difference was found in the numbers of items between the two subjects' basal and ceiling levels. Subject 18's profile covered a range of 21 items spanning a mental age range between 12.4 and 17.8 months; by comparison Subject 19 was presented with 35 items before his ceiling level of 19.7 MA months was reached, his basal level having been determined at 10.5 months. In addition to this, performance differed in almost 35% of the total number of items presented to both subjects. Comparison of the protocols of two further subjects of close but differing ages who also achieved identical scores - Subject 26 (36 months) and Subject 29 (42 months) - revealed a similar diversity of performance, with passes and fails differing in 16 out of 40 (40%) of the total number of items presented. The majority of items on which these two subjects' performance was found to differ fell between their respective ceiling levels. The 36 month old peaked at item 160 which has an age placement of approximately 5 MA months beyond the ceiling level attained by the older subject. These within-profile differences emphasise the need for caution in interpreting this type of developmental index, an issue which will be returned to in Chapter 5.

Analyses of test protocols

A major aim of this study was to investigate whether performance on cognitive tasks was a true reflection of subjects' underlying competence. The data collected is discussed below in separate age groups and in relation to specific subsets of items, but around a common theme: whether, for individual subjects, failure on specific test items was due to inadequate levels of cognitive ability or whether there was reason to believe that inadequate motivation was restricting performance. It will be seen that there were many instances in which subjects' performance did not meet the stringent statistical criteria necessary to score a pass, but that these failures could not be confidently attributed to a straightforward absence of the required level of cognitive ability. Children frequently failed items by default through failure to engage in the tasks presented. Failure to engage was defined as any instance where the child's response fell into one of the three following categories:

1. refusal to attend to the task,

2. rejection of the task object by casting, swiping or dropping it on the floor,

3. repeated production of an inappropriate off-task behaviour in response to the task materials.

It was not possible, for all items, to identify failures by default using these criteria. Items at the bottom end of the scale, for example, are largely concerned with sensory-perceptual acuities and discriminations and the ability to respond appropriately to these (e.g. item 19; turns eyes to ring, item 23: reacts to paper on face, item 32: regards cube, item 47: turns head to sound of bell) and with what is referred to by Lewis et al (1986) as sensory-motor manipulation (e.g. item 43: simple play with rattle, item 44: carries ring to mouth, item 43: manipulates table edge). In these cases there is less scope to distinguish between criterion fails and failures due to the child's failure to engage. This distinction does become increasingly easier to make at higher task levels. There are, however, at all levels in the BSID a small number of 'incidental' items in which there is no particular task to be engaged in (e.g. item 13: vocalises once or twice, item 97: repeats performance laughed at); by definition these too could not be incorporated into the 'failure to engage' analysis.

In addition there were many cases in which subjects failed particular items despite parental reports that their children did in fact possess the requisite skills to pass them. Parental reports were not taken into consideration in the first level quantitative analysis of failures to engage analysis unless it was clear from the child's behaviour that s/he had failed a given item by default. They were, however, used in the more qualitative analyses (see below), particularly in the case of younger subjects.

Table 3:4: Frequency of failures by default on specific items across age groups

Item	Age level: 9	15-21	24	30	36	Total
88: Picks up cup: secures cube	1					1
105: Dangles ring by string				1		1
107: Puts beads in box		1				1
108: Places 1 peg repeatedly		1				1
110: Blue board: places 1 round block		1				1
111: Builds tower of 2 cubes		2				2
114: Puts 9 cubes in cup		2			1	3
117: Shows shoes		2				2
118: Pegs placed in 70 seconds		1	1			2
119: Builds tower of 3 cubes		2			1	4
					(+1 42m)	
120: Pink board: places round block		1				1
121: Blue board: places 2 round blocks		1			(+1 42m)	2
122: Attains toy with stick			1			1
123: Pegs placed in 42 seconds					1	1
124: Names 1 object		1		2		4
					(+1 42m)	

Item	24	30	36	42	48	54	Total
128: Points to parts of doll	1		3	1	1		6
130: Names 1 picture		1	1	2	1		5
131: Finds 2 objects (+1 21m)	2	2		2	1		8
132: Points to 3 pictures			1	2	1		4
133: Broken doll: mends marginally	1	1		1	1		4

Item	Age level 24	30	36	42	48	54	Total
135: Differentiates scribble from stroke			1	1			2
138: Names 2 objects		1					1
141: Names 3 pictures			1				1
143: Builds tower of 6 cubes		1	2	1		1	5
144: Discriminates 2: cup, plate, box		1	2	1		1	5
147: Imitates strokes: vertical and horizontal				1		1	2
148: Points to 7 pictures				1			1
149: Names 5 pictures				1			1
151: Pink board:reversed				1	1		2
152: Discriminates 3: cup,plate,box			2	2	1		5
154: Train of cubes				2	2		4
155: Blue board: completes in 150 seconds				2			2
157: Folds paper			1		1	1	3
158: Understand 2 prepositions				1			1
161: Builds tower of 8 cubes			3				3
Totals	**6**	**10**	**20**	**25**	**10**	**4**	**92**

Total 9m - **1**

Total 15-21m - **16**

Table 3:4 summarises the total number of failures by default observed on individual items by subjects in the age groups, 9 - 54 months (the single 60 month old subject showed no failures to engage). Because only one child was available for participation at 15, 18 and 21 months, these 3 subjects were placed together in a single age group - 15-21 months). The following sections provide a detailed account of the performance/competence differentials found at each age level.

Failures to engage: Analysis by age group

3 months N=8; Subjects 1 - 8
(raw score range 19-33; MA range < 2m - 2.5/3m)

For the reasons outlined above, analysis of failures to engage on items at this lower level is very restricted and is therefore largely made up of parental reports and of observations of subjects' behaviour during testing.

There were few items on which children at this age level could clearly be identified as producing below optimal performances. Most notable was the frequent tendency for subjects to demonstrate a bias to the right side on items which required that they discriminate the direction of a sound source, or that they turn their head towards a visual stimulus. Of the 8 subjects tested at this age, 5 demonstrated this bias - looking towards sounds presented to their right side but not to their left. This bias is most likely due to the tonic neck reflex which is known to be more marked in DS than in other groups of children. One mother, however, described this behaviour as "laziness", and certainly one child who had relatively good muscular control of his head nonetheless showed this bias to one side.

6 months N=5; Subjects 9 - 13
(raw score range 34-62; MA range 2.5/3m -5/5.5m)

Mothers' reports were particularly interesting for this age group - especially those concerned with subjects' responses to sounds. Subjects 9 and 12 both failed item 47 in which the child is required to turn the head towards the sound of a bell rung outside the range of vision. Both were reported to have shown disturbance at loud noises when they were younger, but now to show very little response. Subject 9 was later diagnosed to have a hearing impairment, although he would occasionally react to sounds, albeit very slowly. He was, however, very interested in making noises for himself, persistently banging test objects on the table. Although Subject 12 demonstrated a clear response to sound, she too would not turn her head towards them; in addition to item 47, she also failed item 48 (turns head to sound of rattle). This child's mother reported that she would turn towards her own noise-making toy "because she can do things with it".

A third subject - Subject 10 - failed item 34 in which subjects are required to glance between a bell and a rattle shaken one after the other within their visual field. This subject was reported to turn her head towards sound sources "only when she wants to". Of the two remaining 6 month olds, one was very delayed in turning her head towards the sound of the bell or rattle and would only do this very unreliably, while the other reacted only to the rattle, not the bell which makes a louder sound, indicating that poor response to this item was not due to any hearing impairment.

It is possible that this poor response to sound-making objects at 6 months may be linked to the persistence of the tonic neck reflex seen at 3 months. At that earlier age level several subjects were unable to turn towards the source of a sound on their left. The fact that they did respond to sounds presented from the right side suggests that they may have had the intention to do so, however. It may be that, due to a mismatch between intention and ability at an earlier age, once the ability *does* become available, these children simply do not make full use of it. Cunningham (1979) has offered a similar explanation for the absence of visually directed reaching in many DS infants. He suggests that due to a lack of concordance between visual and motor systems prior to 16-18 weeks, intention and visuo-manual coordination, the necessary prerequisites for the development of reaching, do not emerge in parallel in DS. Although children become motivated to reach toward objects, infrequent success due to lack of motor control results in a general decrease in arm extensions over the pre-reach period. When visual motor coordination does appear, these children then demonstrate more non-functional proximal activity (e.g fingering of the table edge) than reaching, possibly indicating that the visually initiated reaching pattern has been extinguished. In the same way the failure of the 6 month olds in this study to turn towards the source of a sound may indicate that this head turning response has been extinguished.

Many other items at this age level require that children demonstrate a degree of eye-hand coordination and motor control (e.g. item 37: reaches for dangling ring; item 54: picks up cube; item 63 lifts inverted cup). All 5 subjects in this age group had difficulty with at least one such item, but individual subjects responded to this difficulty in different ways. Subjects 12 and 13 persisted in their attempts to pick up and retain the 1 inch cubes, although their responses could not be credited as passes, whereas Subject 9 reached out and touched a cube, but made no attempt to pick it up - despite his mother's report that he was capable of doing so. This same child pushed the inverted cup around the table in an apparent attempt to pick it up but eventually lost interest. Subject 11's first response on this same item was not sufficiently clear to be credited with a pass and on re-presentation she made no attempt whatsoever to pick up the cup.

These different responses may demonstrate two different stages in the development of an attitude of 'learned helplessness' towards grasping objects (see Chapter 1). The lack of success in attaining the cube observed in the first two subjects, despite their clear intention to pick it up, may result in the avoidance behaviour seen in the other two subjects.

There were few items which subjects at this age level failed in common but which, according to mothers' reports, they should have been capable of passing. Subject 9 failed to produce the expected behaviour on 8 items - 16% of the total number of items falling within his basal and ceiling range: had he passed all 8 of these items his MDI would have been 10 points higher than that actually attained. In most cases however, subjects appeared to be underperforming on only two or three items, therefore keeping within the standard error of measurement for this age group.

9 months N=2; Subjects 14 & 15
(raw score range: 79 - 81; MDI/MA range 64 - 69/ 5 - 5.5m)

The 2 subjects tested at this age attained very similar raw scores. Pass and fail patterns were, however, very different: of the total number of items presented to both subjects, there was only 54% agreement. Subject 14 failed to produce the required responses for four items which, according to her mother, she was definitely capable of passing (item 62: looks for fallen spoon; item 69: transfers objects hand to hand; item 84:

listens selectively to familiar words). The latter 2 of these items fall into the 'incidental' category, however, and could not be included in the analysis of failures by default; nor could item 62 as it does not involve a specific task in which the child is required to engage. One item which could be included in the failure by default analysis was item 88 (picks up cup: secures cube). On presentation with this item this subject cast the inverted cup off the table.

Subject 15 also had to have item 88 presented a second time before she would produce the appropriate response. This child in fact had performed so poorly in testing that her mother offered to bring her back for a re-test. Interestingly, although her overall score was improved on this second occasion, she also failed two items which she had passed on the first test.

15 -21 months N=3; Subjects 16 - 18
(raw score range 101-113; MA range 12m - 15m)

Perhaps the most notable feature of performance among the three subjects in this age band was the frequency with which they failed items by default because they cast test objects from the table. Subjects 16 and 18, for example, refused to have anything to do with the red cubes and consequently failed at least 3 items each (items 111 & 119: builds tower of 2 or 3 cubes; items 90, 100 & 114: puts cubes in cup). Other examples of possible failure to perform to full competence were also found. Subject 15 (15m), after having placed one peg on the peg board, completely rejected this task (item 118) while Subject 17 (18 months) refused to engage in an earlier item with this same board (item 87: fingers holes in peg board) after having unsuccessfully attempted to place the pegs. In both cases mothers reported that their children should have been able to demonstrate the required behaviours and indeed both peg board items were successfully passed by both the 9 month olds.

There was also a certain amount of very clear underperformance in evidence. On item 107 in which the child is required to put 6 small beads through a hole in the lid of a box, Subject 18 appeared to be having great difficulty at first, dropping the beads on the lid or around the box. When she was handed the sixth bead however, she placed it very skilfully over the hole and dropped it into the box - almost as if to demonstrate that she *was* capable of passing this item - but only on her own terms. This particular child failed 7 items by default - her MA might have been 2 - 3 months higher and her MDI 16 points higher had she not failed to engage in these 7 items.

24 months N=2; Subjects 19 & 20
(raw scores 119; MDI 54; MA 16m)

As already noted, the 2 subjects tested at 24 months produced very different performance profiles but attained identical scores. Of the 12 items on which performance was found to differ, at least 4 were reportedly within these subjects' repertoires, but were not produced. Both subjects cast several test objects (e.g. the broken doll: item 133; the pink form board; items 120, 137 & 151 and the toy and stick: item 122). Both subjects also produced inappropriate rather than failing responses to other test items. Subject 19, for example, insisted on cuddling the jointed doll when presented with item 128 (points to parts of doll) despite repeated attempts on both her mother's and E's part to persuade her to identify various parts of its body - something which her mother reported she was very capable of doing. Both subjects also refused to cooperate on the item which requires finding an object hidden under one of two cups (item 131), seeming to be deliberately "cheating" by picking up both cups, or by pretending to drink out of them.

30 months N=2; Subjects 21 & 22
(raw score range 117- 134; MA range 16/17m -20/21m)

Again the 2 subjects at this age refused to cooperate on the object concept item just described (item 131). There was little else in common about their performance on specific items, however, as they were presented with very different sets of items, with Subject 21's basal level having been set at the item preceding the final item passed by Subject 22. Scores differed by 17 raw score points - and MAs by around 5-6 months. Despite these differences, in both cases there was evidence of items having been failed by default; Subject 22 in particular failed to produce the required responses to 7 items although she had reportedly been observed to do so previously. Two such items involved verbal responses (item 106:

imitates words and item 113: says 2 words) which were also failed by default by two younger subjects scoring at a similar level to this child. Because these two items are incidental, however, they could not be included in the 'failure to engage' analysis. There may, however, be some developmental connection between failure on these lower-level speech items and the failure to produce higher-level verbal responses seen in the higher scoring child. The mother of Subject 21 reported that her child was very reluctant to name objects on command (items 124, 138 & 146; names 1,2 or 3 objects), but would approach and name objects by herself.

36 months N=4; Subjects 23 - 26
(raw score range 118-141; MA range 16/17m - 23m)

As can be seen, scores varied substantially in this age group. Subject 26's protocol extended from items placed at 14 months to those at the end of the test, at 30+ months. His basal level had to be set at such a low level because he was especially behind in his speech development and consequently unable to pass item 113: says 2 words. He also failed several other items around this age level, however (e.g item 114: puts 9 cubes in cup; item 119: builds tower of 3 cubes), despite passing items at the very upper end of the scale.

In 3 out of the 4 36 month old subjects there was clear evidence of underperformance, both from parental reports and from subjects' responses themselves, which included deliberately dropping or pushing away test objects, flat refusals to engage in the items, or failure to comply with instructions. On average each 3 year old failed approximately 5 items by default - which translates into a difference in MDI of 10 points at 20 months (the median mental age level) and an MA difference of approximately 2 months.

42 months N=4; Subjects 27 - 30
raw score range 126-154; MA range 18-30m

In this age group mental ages also varied widely, with one child (Subject 27) requiring presentation of items at the uppermost age range only, while another (Subject 30) attained a basal level of 14 months. Although this low basal level was set at the item preceding an apparently genuine failure on item 115 (closes round box), several items at the same approximate mental age level were failed by default through casting test objects and failure to engage fully in the tasks.

Three out of 4 subjects of this age produced less than optimal performances on the blue form board task and also refused to attend to the pictures (e.g item 130: names 1 picture; item 148: points to 7 pictures). Failures to engage were less ambiguous among subjects at this stage in development. There were many more examples of flat refusals and failures to cooperate on tasks than at earlier age levels, where it was more frequently necessary to consult parents as to whether failures were genuine - or by default. Again, on average approximately 6 items per child were failed by default - several of which were failed in common by 2 or more subjects (see Table 3:5 below).

48 months N=2; Subjects 31 & 32
(raw score range 139 -152; MA range 23m - 26/27m)

The test performance of both subjects in this group peaked at the same ceiling level - item 161 (builds tower of 8 cubes) - but there was a difference of approximately 6 months in terms of age placement of basal levels. With the exception of speech items, the majority of items in this 6 month range which were failed by Subject 32 (who had the lower basal level) were failed as a result of refusals to attend to test objects or of inappropriate off-task behaviour in response to test materials. In total this child dropped 6 raw score points - accounting for over 2 MA months - through failure to 'perform'. Subject 31 would have attained an MA of 30 months had she not refused to cooperate on 3 of the 15 items between her basal and ceiling levels. Failure to engage in these three items (item 151: pink board: reversed; item 154: train of cubes; item 157: folds paper) - for which she was reported to possess the requisite skills - resulted in an MA score of approximately 3 months lower than she was apparently capable of attaining.

54 months N=2; Subjects 33 & 34
(raw score range 152 - 158; MA range 27m - 30+m)

Again, although both subjects in this age group attained similar ceiling levels, the difference of 8 raw score points between the two scores was due largely to Subjects 34's clear refusal to cooperate on at least 6 items. In comparison Subject 33 failed a total of only 4 items with no evidence that ant of these failures were not genuine.

60 months N=1; Subject 35
(raw score 161; MA 30+m)

This subject proved to be well beyond the mental age level covered by the test. She appeared to be cooperating fully and the two items which she did fail (items 161: builds tower of 8 cubes; 163: understands 3 preopositions) seemed to be due to a genuine inability to produce the required responses.

Failures by default: an overview

Table 3:5 shows the number of cases in which specific types of items were failed by default by large numbers of subjects (all remaining cases of failure by default are summarised in Table 3:4 above). Inspection of this table reveals a group of items which were being avoided in common by subjects toward the mid - upper MA range of the sample. These items have been catergorised according to the types of skills tested (i.e discrimination, crayon and paper skills, cube behaviours etc.). Several items in each category involved the same basic test procedure being used for a series of items of increasing difficulty levels. For example the test situation in which children are presented with the picture cards is broken down into 6 separate items (names 1/3/5; points to 3/5/7). Failures by default were observed at almost every level of this item series; in some cases subjects refused outright to cooperate at any level, in others they would only cooperate up to a point before rejecting the task, although reported as having the ability to pass a higher item in the series (e.g. pointing to only one picture despite reportedly being able to name several).

It can be seen that the most frequently avoided single item - 131 - was that in which subjects are required to find an object hidden under 1 of 2 inverted cups; 8 out of the 9 subjects who failed this item failed it by default. Next came items involving building towers or a train with the cubes, failed by default in 53% of cases of failure. Of items testing discrimination skills, almost 40% were failed through refusal to cooperate in the tasks.

Table 3:5: Frequency of failures by default on specific types of items

	No. Cases failed	No. failures by default	%
Discrimination			
Pictures			
130: Names 1	9	5	55.5
132: Points to 3	9	4	44.4
129: points to 5	6	1	16.6
141: Names 3	7	1	14.2
148: Points to 7	7	1	14.2
149: Names 5	7	1	14.2
Mean	7.5	2.16	26.5
Naming objects (ball, watch, pencil, scissors, cup)			
124: Names 1	8	4	50
138: Names 2	6	1	16.6
146: Names 3	6	-	-
Mean	6.66	1.6	25
128: Points to parts of doll	7	6	85.7
Discriminates cup, plate, box			
144: Discriminates 2	6	5	83.3
152: Discriminates 3	8	5	62.5
Mean	7	5	72.9
Crayon and paper			
125: Imitates crayon stroke	5	-	-
135: Differentiates scribble from stroke	7	2	28.5
147: Imitates strokes: vert. and horiz.	6	2	33.3
157: Folds paper	8	3	37.5
Mean	6.5	1.75	24.8
Cube behaviours			
111: Tower of 2	3	2	66.6
119: Tower of 3	7	4	57.1
143: Tower of 6	8	5	62.5
161: Tower of 8	9	3	33.3
154: Train of cubes	7	4	57.1
Mean	6.8	3.6	55.3

Blue form board			
110: Places 1 round block	2	1	50
121: Places 2 round blocks	2	2	100
129: Places 2 round and 2 square blocks	1	-	-
142: Places 6 blocks	-	-	-
155: Completes in 150 secs.	3	2	66.6
Mean	1.6	1	43.32
Mends broken doll			
133: Marginally	6	4	66.6
131: Finds 2 objects	8	7	87.5

During analysis for the above table, a pattern emerged among the mid - upper MA range of the sample. This pattern indicated possible links between apparently genuine failures on several items by subjects with lower mental age levels, and failures to engage in the same items demonstrated by subjects who attained higher mental age scores. To examine this in more detail the section of the sample attaining MA scores in the mid - upper MA range was divided into 3 groups on the basis of raw scores:

Group 1: raw score range - 113-126 (MA 15-18 months) N=6
Group 2: raw score range - 130-139 (MA 19-22 months) N=6
Group 3: raw score range - 141-158 (MA 23-30 months) N=5.

Item - item default comparisons were then made across these 3 groups.

The link between mental age level and item default seemed particularly strong in items testing discriminatory abilities. Item 128 (points to parts of doll), for example, appeared to have been genuinely failed by 3 subjects in Group 1. Given that this item fell just beyond the ceiling levels established for 3 of these subjects, this was unsurprising. Although this same item fell well within the basal to ceiling ranges of all 6 subjects in Group 2, however, it was failed by 4 subjects in this group - in 3 cases by default. Item 132 (points to 3 pictures) was similarly failed by 4 subjects in the lower group, none of whom were reported to be underperforming; by comparison all 5 failures on the same item Group 2 were by default. Both items 128 and 132 were below the basal levels established for all 5 subjects in Group 3 and were therefore not presented to subjects at this MA level. Item 144 (discriminates 2: cup, plate, box) shows a similar pattern for this higher scoring group, however; failures on this item by subjects in Groups 2 and 3 were largely by default.

This pattern was also evident in items involving the 1 inch cubes: 4 genuine failures to build a tower in the lower MA group contrasted with 2 out of 5, and 2 out of 4 failures being by default in Groups 2 and 3 respectively. Several subjects were also very reluctant to build a train with the cubes: 3 out of 5 failures were by default in Group 2 and 1 out of 2 on the upper MA group.

DISCUSSION

The major aim of this study was to investigate, both quantitatively and qualitatively, the performance of a group of infants with Down's Syndrome on the wide-ranging set of cognitive tasks included in the mental scale of the Bayley Scales of Infant Development.

The quantitative analyses carried out revealed an overall pattern of declining mean DQs with increasing chronological age. This finding that mental age does not rise proportionally with chronological age is consistent with that found in other cross-sectional and longitudinal studies of cognitive development in DS (Oster 1953; Share et al 1961; Koch et al 1963; Loeffler and Smith 1964; Fischler et al 1964; Dicks-Mireaux 1966; 1972, de Coriat et al 1968; Carr 1970; 1975, Eipper and Azen 1978; Ramsay and Piper 1980; Hanson 1981; Schnell 1984; Piper et al 1986; Sharav and Schlomo 1988). Scores were uniformly low; developmental levels at almost every age group were well below those of average non-handicapped children of similar ages, with the gap between the two populations widening over each age interval. Mean raw scores and mental age levels were found to increase steadily however, indicating that DS children do acquire new skills and abilities as they grow older, and that tests such as the Bayley can demonstrate this.

Even in this small sample, however, group means, whether representing DQ, MDI, MA or raw score levels, were frequently found to obscure individual variations. Indeed one advantage of having a relatively small sample is that it does make it possible to see the extent to which individual subjects deviate from the overall pattern. Individual deviation patterns may well be masked in larger studies particularly when analyses are - as is typical - conducted only on group data. The sharp incline of the slope in Figure 3 (group Mental Age levels) at 15 months, for example, clearly reveals that the one child at this age level was performing at an MA level well beyond that of the two children in the two subsequent age groups. This calls into question the value of cross-sectional measurement of a population in which such substantial variation in ability levels can be found (e.g. Carr 1975; Loeffler and Smith 1964; Cornwell and Birch 1969). The mean raw score of 125 calculated from the scores of the two 30 month olds similarly gives little indication of the fact that one of these children was performing at a level equivalent to that of one of the two 42 month old subjects while the other was doing less well than one of the 24 month olds. Nor can this method tell us anything about the expected rate of development of either child; the mean raw score level attained by subjects at 36 months would indicate a predicted *drop* of 3 raw score points by this age for the higher scoring 30 month old.

Even in the case of subjects with highly similar scores, detailed analysis of test profiles revealed that in most cases there was very little agreement in terms of the individual items passed and failed by these children. In many experimental studies requiring children to be pre-matched on overall levels of ability, these differences would generally be overlooked; MA matches are generally made on the basis of scores alone. For example, little account would be taken of the fact that Subject 26 (36 months) and Subject 29 (42 months) whose scores were identical demonstrated very different levels of verbal ability. The younger child who failed almost every verbal item which fell between his basal and ceiling levels was able to compensate for this in terms of score by passing a number of spatial items which fell beyond the ceiling level of the 42 month old. This older child did not demonstrate such a delay in language and consequently produced a narrower, more consistent profile of passes. Can it reasonably be accepted on the basis purely of their scores that these two children are developmentally similar ? More importantly, can it reasonably be accepted that they are both developmentally similar to a younger non-handicapped child who has attained the same score - but have simply reached this developmental stage more slowly?

Difference versus delay

The slow development view of development in handicap is largely justified by its proponents on the grounds that mentally handicapped children achieve the developmental milestones measured in psychometric tests in a sequence similar to that found in non-handicapped children, with the only difference between the two populations being one of rate. Similar scores attained by older handicapped and younger non-handicapped children are offered in support of the view that development in handicap *is* simply slow. On a purely quantitative level, the cross section of scores attained by children in the present study would support this 'slow' theory of development. The analyses which focussed on more qualitative features of performance suggest, however, that there may be important differences between levels of performance as reflected in test scores and true levels of competence in the DS sample tested here. This competence/performance differential is not presumed to exist in theories of normal development or in psychometric techniques. Such findings do however support a growing body of evidence from other studies which suggests that development in children with mental handicap may in fact differ in very fundamental ways from that seen in non-handicapped children (Rogers 1977; Morss 1983,1985; Macpherson 1984; Rondal 1984, 1988; Cherkes-Julkowski et al 1986; Wishart 1986, 1987; Moss and Hogg 1987; Dyer et al 1988).

Intuition alone suggests that a 'slow' theory of cognitive development derived from outcome measures from psychometric tests standardised on non-handicapped children must be inadequate. 'Snapshot' measurements of cognitive ability such as MDI/DQ/IQ performance cannot hope to reflect the very different learning experiences of the two populations. Mentally handicapped children may be slow to reach developmental milestones but this does not mean that their lives whilst approaching these milestones have taken place in slow motion. By nature of the fact that they are handicapped, their experiences of success and failure prior to reaching any given milestone must inevitably have been very different from that of a normal child.

It seems reasonable to assume that an exaggerated experience of failure must at some point affect a child's willingness to persevere in learning attempts. As mentioned previously, Cunningham (1982) suggested that the absence of visually directed reaching in young babies with DS may be an example of repeated failure resulting in the extinction of a motor pattern normally seen in most non-handicapped babies of a similar age. Normal development is structured in such a way that motor abilities and the intention to execute these abilities emerge concordantly. Levels of arousal appear to be sustained by the availability of appropriate motor abilities. A normally developing child can turn towards the source of a sound and therefore learns that it is worthwhile to continue to do so. By comparison, findings from the present study seem to indicate that DS children of the same age lose the intention to locate sounds, possibly as a result of their inability to execute the required motor action at the appropriate stage in development. Ability and intention do merge at a later stage; the 9 month old child with DS can and will demonstrate this ability, but the history of its development is, it would seem, very different from that experienced by the normally developing child.

Similar explanations could be provided for the many examples in this study in which children avoided items which they should have been capable of passing. Past experience of failure in similar situations may simply have reduced these children's motivation to engage in tasks requiring demonstration of skills that had had particularly 'difficult' learning histories. One such skill may be the processing of auditory information. DS children have been found to characteristically demonstrate a short term memory deficit for information presented aurally (Bilovsky and Share 1965; Burr and Rohr 1978; McDade and Adler 1980; Marcel and Armstrong 1982). In this study items which involved verbal instruction (e.g. item 128: points to parts of doll) or discrimination (e.g. between a house and three other objects displayed on a picture card: items 130, 132, 139, 141, 148, 149) were beyond the overall ability levels of children in the lowest of the three groups (group 1) selected for the purpose of demonstrating links between genuine failures and failures to engage at differing levels of ability. It could not be concluded therefore that there was any such deficit or lag apparent among children at this stage of development. Although children in group 2 were reported to have acquired the requisite skills to pass these same items which, in all 6 cases, fell toward the lower to middle ends of their performance profiles, they almost universally refused to demonstrate them. Children in this group were clearly well in advance of children in group 1 on other skills tested by items around the same MA level as those involving discriminatory abilities. There was far less evidence of deliberate underperformance on the blue form board items, for example. The high frequency of failures by default observed in the same children on the discrimination tasks would seem to suggest that this skill had had a comparatively more difficult learning history. This interpretation was further supported by the large number of below optimal performances observed by children in groups 2 and 3 on higher level discrimination tasks (items 144/152: discriminates 2/3: cup, plate, box)

Obviously no conclusions about the developmental role of different learning histories can be drawn from this - or any - cross-sectional study. Longitudinal investigations of the performance of individual children on these types of items would be necessary to determine the origins of such patterns. What can be concluded from this study, however, is that although the acquisition of discriminatory abilities may be delayed in relation to the development of other skills, the fact that DS children are reluctant to demonstrate this ability even once acquired suggests that it is insufficient simply to explain the development of this skill in terms of a deficit or a delay in relation to normal development.

It may also be the case that to describe development in handicap as simply being different is equally insufficient. Indeed a great many studies comparing aspects of development in DS and non-handicapped children have failed to produce evidence of significant differences between the two populations (e.g. Cunningham et al 1985; Breeghly 1987; Pruess et al 1987; Metcalf and Stratford 1986). In a detailed study of the language development of children with DS, Rondal (1988), moreover, was able to provide examples *both* of differences *and* of apparently straightforward delays. He argues that there is no clear dichotomy between difference and delay, but that the distinction is dependent on the level of analysis. In relation to language development, he maintains that a solution to the debate depends largely on whether differences are being sought within the system of language as a whole, or in its separate components. If there is a delay in varying degrees for different aspects of language, development within these separate aspects can simply be described as being delayed, but for language as a whole, development must be seen as different.

The development of the 36 month old DS child in this study who demonstrated a specific delay in language acquisition can similarly be seen either as delayed with respect to that particular aspect of cognitive development, or as different in terms of the structure of development as a whole. It is well established that the development of language in DS children commonly lags behind that of other abilities. Many proponents of the 'difference' theory of development use this observation to support their position. By comparison, however, the similar scoring 42 month old did not demonstrate a delay in language relative to her overall level of performance. The more concentrated pattern of passes on items within one particular mental age range could be taken to indicate that in this particular case there was no evidence that development in DS differs from the 'normal' pattern; this subject's poor score was simply symptomatic of delayed progress through the normal sequence of developmental milestones.

It would in fact appear that the test profile from either of these two children can tell us very little about differences and delays in DS development. Failure to produce the verbal responses required to be credited on an assessment test does not necessarily indicate a delay in language acquisition. On the basis of both their behaviour and parental reports both children were in fact producing less than optimal performances, thereby rendering invalid any conclusions regarding the relationship between specific and overall levels of ability. Adoption of a level of analysis based purely on what children overtly demonstrate themselves to be capable of may not be sufficient to distinguish between developmental delay or difference in DS.

Motivational differences: assessment and handicap

Many proponents of either delay or difference theories fail to take into account the influence that motivational factors can have on test performance and on cognitive development. One way in which DS children do appear to be clearly different from normally developing children is in their reluctance to demonstrate behaviours which are within their repertoires. It is stated in the test manual that the items selected for inclusion in the BSID were those found to be the most effective at eliciting demonstration of specific abilities from the standardisation sample of non-handicapped children. From this study it would appear that they are considerably less effective in doing so with handicapped children. In itself this could be offered as evidence that development in handicap differs from the 'norm'.

Whether demonstrating developmental difference or delay, it is very clear from the frequency with which children were avoiding BSID items that they reportedly should have passed that this test is not an effective means of accurately assessing the cognitive abilities of children with DS. Despite this it is still very widely used for this purpose both in clinical settings and in research. Mental age matches made on the basis of outcome measures from the BSID and from other, similar tests of infant development are still common in research studies comparing handicapped groups of differing aetiologies or groups of handicapped and non-handicapped children. Paradoxically, a large number of such comparative studies are carried out for the purposes of demonstrating developmental differences between the populations in question (e.g Marcel et al 1988; Mundy et al 1988; Loveland 1987; Cardoso-Martins et al 1985). The BSID in particular is still widely acclaimed as being "extremely well standardised", "very reliable" and "unrivaled at determining a child's developmental status relative to its age mates (Francis, Self and Horowitz 1987). In a recent international study of the performance of DS children on the BSID, the authors concluded that "with DS children Bayley tests can and should be used beyond the stage of 30 months as long as the test captures the performance of the child" (Rauh and Rudinger 1987). The study presented here has shown that for DS children, whether below or beyond 30 months, full levels of ability are not captured in BSID performance.

Results from assessments made in infancy are very often consulted in making decisions regarding school placements and eligibility for other services provided for the handicapped. The present study suggests strongly that single assessment sessions do not provide an accurate picture of the developmental status of chldren with DS and indicates that such measures should be regarded with more caution than would currently appear to be the case. In many cases, subjects were seen to be underperforming by up to two months below their optimal levels of performance. Given that children toward the upper end of the sample were performing at mental age levels of almost half their chronological ages, this difference is not insubstantial.

Despite the wide dependance on scores from tests such as the BSID, few studies have directly investigated their reliability in children with mental handicap; testing is generally carried out only once and typically by a stranger to the child. One exception is a study carried out with groups of handicapped, non-handicapped and minority children by Fuchs and Fuchs (1985) who found that on re-testing, considerable improvements in scores were achieved; this was attributed by Fuchs and Fuchs to examiner familiarity. Given that in the present study, behaviours witheld in testing had been observed by individuals more familiar with the children, it is possible that a more accurate picture of ability levels could be obtained by repeating the tests with a familiar examiner. The next chapter will therefore present the results of a study in which DS children were assessed and re-assessed by the same examiner at closely spaced intervals.

CHAPTER 4

THE PERFORMANCE OF DS CHILDREN ON THE BSID: RELIABILITY AND VALIDITY

The quantitative analyses presented in the previous chapter of the BSID performance of a group of infants and young children with DS revealed that levels of ability, as reflected in test scores, were consistent with those found in other cross-sectional studies of the cognitive development of this population of children. Detailed analyses of children's performance and parental reports indicated, however, that in many cases, demonstrated levels of performance fell short of previously demonstrated competence. An obvious implication of this finding is therefore that scores were not accurate indices of children's optimal levels of ability. This in turn lends support to an argument offered by many parents of children with mental handicap: that current test procedures are not adequate means of assessing the true cognitive abilities of their children.

Indeed a constant source of frustration to parents is the frequency with which their children do tend to underperform in assessment situations. Perhaps more frustrating, however, is the fact that although this has not gone unnoticed by professionals there seem to have been few attempts to investigate this problem in detail. This seems surprising given the frequency with which these children are formally assessed in comparison with non-handicapped children and the importance of the decisions made on the basis of these assessments.

During testing for the previous study there were many instances of children failing items for which they were known to have acquired the requisite skills and to have demonstrated these skills in other settings. From this it could be suggested either that the performance of these children in assessment situations is not sufficiently reliable to justify dependence on single testing sessions, or that the test procedure itself is not a valid means of assessing their cognitive development. The present study aims to explore these issues in detail.

It has already been mentioned that the question of reliability has been tackled indirectly by a group of researchers who had pinpointed examiner unfamiliarity to be a major cause of underperformance in handicapped pre-school children in single assessment situations (Fuchs and Fuchs 1985). They pointed out that in the majority of clinical and educational settings, examiners are strangers to the children they test. This was substantiated both by reports from practicing professionals and by an analysis of the user manuals of 20 well known intelligence and speech/language measures; in only two of these manuals was there any recommendation that examiners establish pretest contact with their examinees. Fuchs and Fuchs' review of a large number of investigations of the effects of examiner unfamiliarity on test scores revealed that this factor particularly compromised the reliability of scores attained by handicapped children: the majority of studies in which children performed significantly more poorly with unfamiliar examiners were found to involve handicapped, minority and/or pre-school children. In addition, this review was able to show that the test performance of handicapped pre-schoolers was particularly adversely affected by examiner unfamiliarity. Their own follow-up study directly comparing the effects of unfamiliar testers on handicapped and non-handicapped preschool and school-age children found that only the groups of handicapped children performed differentially with familiar and unfamiliar testers (Fuchs et al 1985; Fuchs and Fuchs 1985).

The examiner unfamiliarity hypothesis is not entirely incompatible with the motivational hypothesis which is central to this thesis. The interaction of both variables - examiner unfamiliarity and motivation - might explain the high level of discrepancy recorded between parental reports of ability and DS children's performance on the BSID in Chapter 3. Children had - when in the presence of a familiar adult (i.e. a parent) - demonstrated that the skills necessary for passing certain items were in fact within their repertoires. In testing, many such items were avoided, however, and it was suggested in Chapter 3 that this avoidance might have been attributable to childrens' past experience of failure in similar situations. If this explanation can be accepted, it does not seem unlikely that a direction from an unfamiliar adult to produce a skill which has had a particularly difficult learning history should be met with a certain amount of resistance on the part of the child.

However, although capable perhaps of explaining a proportion of cases in which children were underperforming, it cannot be assumed that examiner unfamiliarity is responsible for all instances in which handicapped children have failed to demonstrate optimal levels of ability in testing situations. Fuchs et al were able to demonstrate that single testing sessions with unfamiliar examiners were not sufficient to attain reliable measures of the abilities of their subjects, but they did not deal with the issue of reliability in any detail. Their analyses were conducted on scores only; no reference was made to the reliability of individual test items. Hence, although scores were found to improve when children were tested with a familiar examiner, they could give no indication as to whether these higher scores reflected a straightforward improvement in performance on a small number of specific items or whether, despite achieving higher overall scores, subjects might not have reproduced all passes attained in previous sessions. Moreover there were no reports of qualitative aspects of children's performance in either session - whether involving familiar or unfamiliar examiners. In the absence of this sort of data it could not therefore be assumed that these improved scores necessarily reflected optimal ability levels.

In this study the question of reliability will be tackled more directly. Children will be tested on two closely-spaced occasions by an examiner who is already familiar with them. In addition to straightforward score comparisons, test protocols will be compared for evidence of changes in performance on specific items. More qualitative aspects of children's behaviour during testing will also be recorded.

The second question to be addressed in this study is that of the validity of tests such as the BSID with mentally handicapped subjects. The suggestion was made in the previous chapter that although the items selected for inclusion in tests such as the BSID had been demonstrated to be the most effective means of eliciting evidence from non-handicapped children that they had acquired specific skills, it is not necessarily the case that they are equally effective in doing so with children with mental handicap. At a recent conference on mental handicap the father of a child with DS objected to current assessment procedures on the basis that they require children "to perform like dancing bears". This description would seem to be especially applicable in cases where handicapped children are tested on items which have been designed for use with much younger non-handicapped children. The pattern of scores presented in Chapter 3 clearly demonstrates that as DS children grow older, the gap between their chronological and mental ages increases. Only one subject was performing at the level which would be expected from a similarly aged non-handicapped child. This difference in chronological age between handicapped and non-handicapped children of similar levels of cognitive ability is simply overlooked in 'IQ' type tests where the same sorts of tasks are used for testing the abilities of the two populations of children. Failures on tasks are interpreted, for both populations, as evidence that the skills they are designed to test for have yet to be acquired. When an older handicapped child is seen to fail items which are generally passed by much younger children it is generally assumed that his/her development is 'slow'. In the absence of performance data on similarly aged non-handicapped children on these sorts of tasks, such conclusions cannot be justified. The tasks may simply - as frequently appears to have been the case with children in the cross-sectional study - have failed to engage the older child: again, a case of "won't do", rather than "can't do" (Koegel and Mentis 1985, Wishart 1987).

This is not to imply that all cases of failure in test situations are by default and cannot therefore be attributed to deficits in cognitive ability. There were many examples in the previous study of children showing genuine difficulties with specific BSID tasks. It will, however, be recalled that DS children in the errorless learning study (Chapter 2) were more likely to perform poorly following presentation with a task which they perceived to be difficult, but that non-handicapped children did not demonstrate this differential response. Many parents of mentally handicapped children express concern that such differences in perceptions adversely affect their child's performance in a test situation. Whereas a non-handicapped child is likely to perceive a testing situation as an opportunity to demonstrate what s/he *can* do, for a child with mental handicap the same situation is often one in which s/he has little option but to concede to what s/he cannot do.

It seems unlikely, given the substantial and inevitable differences in learning histories, both in terms of length and experience of success/failure, that handicapped and non-handicapped children should approach assessment situations in similar ways. The BSID, like most other tests of this sort, included only non-handicapped children in its standardisation sample. Issues such as examiner unfamiliarity, suitability of test items, and lowered expectations of success were given little consideration. As the BSID is one of tests most widely used with young mentally handicapped children, it seems important to examine the assumption that its reliability and validity are unaffected when used with this population. The major aim of this study was therefore to investigate the reliability of performance of infants and young children with DS on the BSID by comparing performance on two closely-spaced testing sessions administered by the same examiner. The issue of validity was also addressed.

METHOD

Subjects

Eighteen DS children in the age range 6 months to 4 years took part in this study, three at each of the following ages 6, 12, 18, 24, 36 and 48 months. All subjects had also participated in the study reported in Chapter 3.

Procedure

All children were tested with the BSID on two occasions. Testing sessions were separated by one week at the two younger age levels and by two weeks at older ages. The same examiner was used in both sessions. All subjects were already familiar with the examiner from previous visits to the laboratory. Procedure was identical in both sessions and was in accordance with that described in the previous study.

RESULTS

As in the cross-sectional study both qualitative and quantitative analyses of BSID performance were carried out.[1]

Raw scores

Table 4:1 shows the mean raw scores and corresponding MAs in the two sessions for the six age groups tested. As in the cross-sectional study, it can be seen that although raw scores increased significantly with increasing age, even the six month olds scored 1 to 2 months below the norm for their age, with this deficit in performance widening as age increased.

Table 4.1: Raw Scores and Mental Ages

Age (months)	**Testing I**	**Session II**
	Mean Raw Scores (Mental Age)	
6	55. 3 (4.5)	53.3 (4.5)
12	77.3 (7.0)	75.6 (7.0)
18	100 (11.0)	99.7 (11.0)
24	109 (13.6)	108.7 (13.0)
36	131 (19.0)	130.3 (19.0)
48	142 (22.0)	142.3 (22.0)

Reliability of performance

It can also be seen from Table 4:1 that mean raw scores do not differ greatly over the two sessions; gains and losses in scores amounted to a maximum of two points and these had no effect on MAs. A t-test comparison of raw scores indicated that there was no significant

[1]These findings have already been reported in The British Journal of Educational Psychology (J.G. Wishart and L. Duffy, (1990)) Instability of Performance on Cognitive Tests in Infants and Young Children with Down's Syndrome, 60, 10-22.

difference in performance over the two sessions (t = 0.76, (df 17); NS). The correlation coefficient for the two sets of scores was 0.98.

A second measure of reliability of performance, however, produced a very different picture. Examination of individual performance profiles showed that even where raw scores were very similar over the two sessions, performance had in fact changed over a substantial number of items. Subject 14 (36 months), for example, attained a raw score of 128 (MA 19m) on both occasions. Her performance was found to vary on 6 out of 23 - or 26% of the total number of items presented over the two sessions, however. Similarly, whereas a comparison of raw scores achieved by Subject 12 (24 months) indicated a minor drop in performance in the second session (3 raw score points), item-item comparison revealed that his performance had in fact differed on 11 items - 31% of the total of 35 items presented. Half of the children in this sample in fact performed differently on 25% or more of items presented. Table 4:2 summarises the number of items on which performance varied over the two sessions for each age group.

Table 4:2 Performance Variability

Age (Months)	Total No. Items Varying over sessions		Total No. of Items Presented	% Agreement
	Fail to Pass	Pass to Fail		
6	12	19	130	76.2
12	10	15	146	82.9
18	11	13	86	72.1
24	16	16	108	70.4
36	12	25	97	72.2
48	13	13	113	77.0
Total	**74**	**91**		**Mean 75.1**

Although item-item agreement over sessions in individual children would seem to be the most obvious and sensitive measure of a test's reliability, no such normative data are provided with the BSID. The only measure of reliability presented comes from a test-retest study which used the 1958-60 version of the Scales (the immediate predecessor of the current version). The sample on which the current version of the BSID was founded involved 1262 infants. In contrast the reliability study involved only a very small sample of children (28) of only one age level (8 months) and with only the reliability of a restricted number (59) of test items reported (Werner and Bayley 1966). Item-item agreement figures were only reported for items showing the highest and lowest reliabilities; no reference was made to the remaining 17 items which were in the mid-range. Moreover of the 21 items which did obtain particularly low reliabilities, only 11 appear in the present version of the test. It was not therefore possible to make an overall comparison of reliability figures obtained for all of the items included in the present study and those given by Werner and Bayley. A t-test comparison of the test-retest agreement figures obtained for those items showing the highest reliability in the Werner and Bayley study with reliability found for the same items in the present study, however, revealed a significant difference in favour of the Werner and Bayley study ($t = 3.71$, df 21, $p < 0.005$) indicating that DS performance was less reliable even on items which normally produce highly reliable performances.

The only other data on reliability of the BSID comes from an item-item reliability study by Horner (1980) of the test-retest performance of non-handicapped children at two ages, 9 and 15 months. Only 76 items were presented in the Horner study; items 70 - 109 were presented to 9 month old children and items 100 - 145 to 15 month olds. Mean percentage agreement for the entire range of 76 items was 84.55%. This compares with a reliability figure of 75.3% for the same items in the DS group. A t-test comparison of DS percentage agreement figures with Horner's data produced a similar result to the comparison with Bayley's reliability figures ($t = 4.267$, df 75, $p < 0.0005$), that is, DS performance was again found to be less reliable.

Although no indication of individual children's test-retest reliability is provided for the BSID, what is provided is the standard error of measurement at each age level. This allows variability in performance in each individual DS child's scores over the two testing sessions to be compared to the variability which would be expected on the basis of the

SEm for the range of mental ages covered by the test. The SEm was, however, calculated from the distribution of MDI scores and, as will be recalled, most DS scores were too low to be converted into MDIs. Comparisons were therefore carried out between the variability predicted in raw scores on the basis of the SEm and MA level observed for each child. For example: the performance of Subject 12 (24 months) was found to vary on 11 items. This subject scored at an MA level of 16 months. The SEm for 15 months (the nearest age level for which an SEm is provided in the BSID) is 6.7 which, using the norms tables, translates into a maximum of 4 raw score points at this mental age level. Hence, using the SEm, a raw score obtained at this level would be expected to fall within 4 points of the 'true' raw score. Clearly, the variability in Subject 12's performance was in excess of that expected on the basis of the SEm. A t-test between observed and expected variability in all subjects indicated that there was a significant difference over the two sessions ($t = 2.78$, df (17), $p < 0.01$) again indicating that DS performance was less reliable than that of non-handicapped children.

Direction of performance variability

Remaining with the example of Subject 12, of the 11 items on which his performance was found to vary, 7 changed in a pass-to-fail direction, (i.e. were passed in the first session and failed in the second) and 4 in the opposite direction. Although there were two cases in which performance varied in only the pass-to fail direction, most children, like subject 12, showed a combination of both passes-to-fails and fails-to-passes. In total there were 165 cases of unreliability: in 74 cases, items which had been failed in the first session were passed in the second session (fail-to-pass); in 91 cases the change was in the pass-to-fail direction (see Table 4:2).

A fail-to-pass change could result from a genuine developmental advance having been made between testing sessions. Given, however, that sessions were separated by a maximum of two weeks, this was unlikely to be the case when performance was found to improve on as many as 8 or 9 items in the second session. As the examiner was already familiar with subjects on the first testing session, it is also unlikely that increased examiner familiarity could have been responsible for such improvements. More likely is that fluctuating levels of task engagement were resposible for the unreliability shown. In this respect a pass-to-fail pattern is less ambiguous. Clearly the required behaviour was in the child's repertoire but was not for some reason being produced in the second testing.

Since Bayley items in general cannot be passed by chance, any passes, whether produced in the first or second session, are likely to have been genuine. An optimal score was therefore calculated for each child in which credit was given for all the items on which a pass had been scored in either session. Not surprisingly, this score proved to be consistently higher than the score actually achieved in either testing session (1st/2nd session: $t = 2.60/3.24$ (df 17); $p < 0.01/0.005$).

Difficulty level of unreliable items

An analysis of the difficulty level of those items which proved to be unreliable was carried out by relating each item which varied to the developmental level of each child. The midpoint between lowest basal and highest ceiling levels attained over the two sessions for each child and the numbers of pass-to fail and fail-to-pass items which fell above and below this midpoint were calculated. Numbers of fail-to-pass items fell equally on either side of this midpoint. There was, however, a trend for more pass-to-fail items to occur above than below this midline and for these to exceed the number of fail-to-pass items above this point ($t = 1.40/1.24$ (df 17); $p < 0.10$/NS respectively).

Items showing poor reliability

Table 4:3 shows those test items on which performance was found to vary most over the two sessions in the 6 age groups, together with the test-retest item agreement figures available from Werner and Bayley (1965), and from Horner (1980). As indicated by the t-test comparisons presented above, the variability in DS performance was greater than would have been expected on the basis of either of these two reliability studies.

Table 4:3 Items Showing Poorest Reliability

Item	No. Instances of unreliable performance						No. of S's*	Test-Retest % Agreement			
	Age Group							DS subjects	Non-handicapped subjects		
	6	12	18	24	36	48			Horner 9m	Horner 15m	Werner & Bayley 8m
45: Inspects own hands	1	1	-	-	-	-	4	50	-	-	-
53: Mirror image approach	2	-	-	-	-	-	4	50	-	-	-
55: Vocalises attitudes	3	1	-	-	-	-	4	0	-	-	-
56: Retains 2 cubes	1	1	-	-	-	-	4	50	-	-	-
60: Reaches persistently	1	1	-	-	-	-	6	66.6	-	-	-
62: Turns head after fallen spoon	2	1	-	-	-	-	6	50	-	-	-
67: Sustained inspection of ring	2	-	-	-	-	-	5	60	-	-	-
70: Picks up cube deftly and directly	1	1	-	-	-	-	5	60	100	-	91
77: Retains 2/3 cubes offered	1	2	1	1	-	-	6	16.6	92	-	95
81: Cooperates in games	-	1	1	1	-	-	7	57	71	-	58
82: Attempts to secure 3 cubes	-	1	2	1	-	-	7	43	71	-	-
86: Uncovers toy	-	1	1	-	-	-	6	66.6	92	-	85
88: Picks up cup: secures cube	-	-	2	1	-	-	6	50	92	-	-
98: Holds crayon adaptively	-	-	1	2	-	-	8	63	100	-	92
99: Pushes car along	-	-	2	1	-	-	8	63	96	-	-
104: Pats whistle doll in imitation	-	1	-	2	-	-	7	57	62	70	-
106: Imitates words	-	-	1	1	-	-	5	60	88	71	88
107: Puts beads in box	-	-	1	2	-	-	5	40	96	96	-
108: Places 1 peg repeatedly	-	-	1	1	-	-	4	50	92	100	-

Item	No. Instances of unreliable performance — Age Group 6	12	18	24	36	48	No. of S's*	Test-Retest % Agreement — DS subjects	Non-handicapped subjects — Horner 9m	Horner 15m	Werner Bayley 8m
111: Builds tower of 2 cubes	-	-	1	1	-	-	4	50	63	-	-
112: Spontaneous scribble	-	-	1	1	-	-	5	60	-	75	-
113: Says 2 words	-	-	-	1	-	1	6	66.6	-	75	-
114: Puts 9 cubes in cup	-	-	-	2	1	-	7	57	-	75	-
115: Closes round box	-	-	1	1	2	-	7	43	-	92	-
117: Shows shoes	-	-	1	2	2	-	7	29	-	92	-
122: Attains toy with stick	-	-	-	2	2	1	6	16.6	-	67	-
125: Imitates crayon stroke	-	-	-	2	1	1	7	43	-	75	-
133: Broken doll: mends marginally	-	-	-	-	-	2	6	66.6	-	79	-
134: Pegs placed in 30 secs	-	-	-	-	1	1	6	66.6	-	75	-
135: Differentiates scribble from stroke	-	-	-	-	1	1	6	66.6	-	83	-
140: Broken doll mends approx.	-	-	-	-	1	1	4	50	-	67	-
143: Builds tower of 6 cubes	-	-	-	-	1	1	4	50	-	96	-
151: Pink board: reversed	-	-	-	-	-	2	4	50	-	-	-
161: Builds tower of 8 cubes	-	-	-	-	-	1	2	50	-	-	-
162: Concept of one	-	-	-	-	-	1	2	50	-	-	-

* *This column represents the Number of Subjects who were presented with these items*

Many of the items on which performance was found to vary most were the same items on which children in the cross-sectional study had failed by default due to failure to engage fully in the tasks. This link was particularly strong for the 'tower of cubes' items (items 111/119/143/161; builds tower of 2/3/6/8 cubes). 53% of children who failed items in this series failed them by default in the cross-sectional study; here, 41% of children presented with these items changed performance over the two sessions. Other particularly unreliable items which large numbers of children failed by default in the previous study included those involving the crayon and paper (item 125; imitates crayon stroke, item 135; differentiates scribble from stroke) and the broken doll (items 133/140/153; mends broken doll marginally/approximately/exactly). These again proved to be particularly unreliable in this study.

Failures to engage

It became very apparent during testing that in many cases where passes were not being repeated on the second session, or children were passing items they had failed in the first session that failure had been by default. For example, subject 17 (48 months) very capably and willingly pointed to 3 of the pictures in item 141 in the first session, but simply refused to cooperate on the same item in the retest. Similarly, he would not even attempt the crayon and paper items in the second session, and actually broke the crayon in a fit of rage, despite having passed both items 125 and 135 (imitates crayon stroke and differentiates scribble from stroke) in the first session. In contrast he was far more willing to attempt item 151 (pink form board: reversed) in the re-test, actually succeeding in passing it on this occasion, whereas he had repeatedly insisted on turning the form board to its original position before he would place any of the three blocks when presented with this item in the first session.

Given the high incidence of failures by default observed over both sessions, it seemed worth investigating the extent to which variable performance could directly be attributed to children failing to engage in the tasks on either of the two testing sessions. Forty seven out of the 165 cases of unreliability had to be excluded from this analysis because it was not possible for these items to distinguish between criterion fails and failures by default (see p. 55). It is perhaps worth noting however that as

Table 4:4 Items in which subjects most frequently failed to engage

Item	No. Subjects failing to engage						Subjects showing unreliable performance	
	12	18	24	36	48	Total	No.	% Overlap
77: Retains 2/3 cubes offered	2	1	1	-	-	4	5	80
82: Attempts to secure 3 cubes	1	1	1	-	-	3	4	75
88: Picks up cup: secures cube	-	2	1	-	-	3	3	100
18: Holds crayon adaptively	-	1	2	-	-	3	3	100
100: Puts 3 or more cubes in cup	-	1	1	-	-	2	2	100
104: Pats whistle doll in imitation	1	-	2	-	-	3	3	100
107: Puts beads in box	-	1	2	-	-	3	3	100
111: Builds tower of 3:cubes	-	1	1	-	-	2	2	100
114: Puts 9 cubes in cup	-	-	2	1	-	3	3	100
115: Closes round box	-	1	1	2	-	4	4	100
117: Shows shoes	-	1	2	2	-	5	5	100
125: Imitates crayon stroke	-	-	1	-	1	2	4	50
130: Names 2: picture	-	-	1	1	-	2	2	100
135: Differentiates scribble from stroke	-	-	-	1	1	2	2	100

performance on these items did not in all cases change in the fail-to pass direction, it seems highly probable that something other than lack of the required cognitive ability was determining poor performance in one or other session.

In the 118 remaining eligible cases subject records were re-examined for evidence of failures which met the criteria for failures to engage given in Chapter 3. In over half the cases, failure was clearly due to the child refusing to engage in the task: in 40/64 pass-to-fail cases and in 27/54 fail-to-pass cases.

Table 4:4 shows the items which were most often failed through children failing to engage in the tasks. Comparison with Table 4:3 reveals that, to a significant extent, performance variability can be attributed to failures by default.

The difficulty level of each item failed by default by individual children was calculated using the basal - ceiling midpoints. Instances of failure to engage were found to occur equally as often above (32) as below (34) these midpoints indicating that subjects were as likely to avoid relatively easy items as they were to avoid more difficult items.

DISCUSSION

On a strictly quantitative level, results from this study again lend support to the 'slow' development theory. On average, performance in the DS subjects was at approximately half the level expected from non-handicapped children of similar chronological ages. Comparison of individuals' test-retest profiles, however, once again indicated that the performance of this group of DS children did not reliably reflect their 'true' levels of competence. This held true even after allowing for the variability expected between test and retest scores on this particular test.

It will be recalled that the practice of MA matching children with DS with much younger non-handicapped subjects was avoided in the study presented in Chapter 2 on the basis that it implicitly prejudges the delay/difference issue. The validity of this practice was also questioned on the grounds that two children with widely differing performance profiles can achieve identical MAs. Evidence in support of this objection was presented in the DS cross-sectional study in Chapter 3. The results of the present study demonstrate that not only can two different DS children attain similar scores via different routes, but also that it is possible for the same DS child to produce two very different performance profiles over closely-spaced testing sessions which nevertheless result in two identical, or very similar, scores. It will be recalled that no significant difference was found between scores attained in the two sessions. The reliability coefficient found - 0.98 - was in fact substantially higher than typically reported from other tests of cognitive ability in infancy (Herring 1937, Hindlay 1960, Gilliland 1948, Conger 1930). Clearly comparisons of this sort - made purely on the basis of scores and giving no indication of how scores on either session were attained - are not adequate measures of test reliability.

Comparison with normative studies

Reliability coefficients were not reported for overall scores in the Werner and Bayley study. As already noted, only percentages of test-retest agreement were presented for those items administered to their reliability sample of 8 month old infants, this ranged from 41% - 95%. Despite substantial variations in performance on individual items, the overall performance of DS children came very close to meeting Bayley's mean for test-retest reliability: the mean agreement across the two examinations in Bayley's own study was 76.4%, while in the present study the mean figure obtained across the 6 age groups was 75.1%.

Bayley's mean reliability figure of 76.4% is very low and has been given surprisingly little attention either in research studies or in educational practice. The Bayley is generally regarded to be a highly reliable test and is still in wide use with both normal and handicapped infants 30 years after its first appearance. There are a number of reasons to suggest, however, that the 76.4% figure quoted in the BSID manual is a poor indicator of the item-item reliability for the current version of the test. It is, for example, implied in the manual that the overall figure of 76.4% was 'pulled down' by the high level of variation observed in the Werner and Bayley study on items concerned with performance of a social-interpersonal nature. Whereas test-retest reliabilities for the 21 items falling into this category ranged between 41-74%, a range of 85-95%

was obtained for the most reliable third of the 59 items used in that study. The Werner and Bayley study, however, was carried out on the 1958-60 version of the Scales - the immediate predecessor of the current version. Although this is fleetingly mentioned in the BSID manual there is no reference to the fact that almost 50% (10/21) of the social-interpersonal items on which performance was found to be most variable have in fact been excluded from the 1969 version of the test. This would suggest that the overall test-retest figure for the current version must in fact be higher than 76.4%.

This suggestion is supported by the contradiction found between the outcome of the above comparison and that of the SEm comparison. Although the DS mean percentage agreement figure is similar to that obtained in the Werner and Bayley study, when individual DS subjects' reliability figures are compared with the SEm this results in a significant difference in favour of the standardisation sample. Using this method of comparison DS performance was found to be considerably more variable. Again this result indicates a discrepancy between the 76.4% agreement figure obtained in the Werner and Bayley study and the 'true' reliability figure for the current version of the test.

Further evidence of much higher levels of overall and item-item reliability come from the recent study by Horner (1980) in which the average agreement figure was 84.55% and item-item reliability varied between 54% and 100%. Of the two studies, Horner's is preferable as an estimate of the test-retest reliability of the BSID for a number of reasons. Firstly, it has the advantage that it was carried out more recently and used the current version of the test. Secondly, it involved a somewhat less restricted sample of children: 24 infants at both 9 and 15 months in contrast to Werner and Bayley's sample of 28 infants at only one age level (8 months). Thirdly, and most importantly, the test was presented in accordance with the standard testing procedure. Although it is not entirely clear, it is implied in the Werner and Bayley study that all children were given the same number of items (59), irrespective of their ability levels. No such procedural variation is evident in the Horner study which, because it included two age groups, has the additional advantage of including a substantially larger number of items (76).

A further reliability comparison was therefore carried out on the DS data by focussing on the same group of 76 items used in Horner's study. DS mean performance was found to be considerably more variable than in Horner's sample of non-handicapped children (DS mean percentage agreement 74.3%; Horner study 84.55%).

Unreliable performance: the examiner unfamiliarity hypothesis

Subjects in the Horner study were tested once in a clinic setting and once at home. Performance in the younger age group was found to be influenced to a relatively greater extent by the setting, cumulative experience with the examiner and test item familiarity. These variables were found to influence the distribution of passes and fails on inconsistently performed items with there being a statistically significant greater likelihood that the failure would occur in the clinic when it was also the first testing. This pattern of variability did not however affect overall scores, a finding which is at variance both with those of Fuchs and Fuchs (1985) and of Durham and Black (1978) who found that there was moderate improvement in scores among 16-21 month olds when home testing followed testing in an unfamiliar setting. Despite this, an overall similarity between the 3 studies is in the direction of performance variability: where setting or familiarity with the examiner or with test items was found to affect performance on individual items this was predominantly in a fail-to-pass direction. Concordant with findings from the Horner study, DS children's scores in the present study did not significantly differ over sessions. Variability in performance in individual test items was, however, as likely to lie in a pass-to-fail as in a fail-to-pass direction. This result lends little support therefore to the hypothesis that DS children will perform better if they are more familiar with the examiner, the test items and the test setting.

Indeed, given that there was a greater incidence of pass-to-fails than fail-to-passes it could be suggested that increased familiarity with the test setting actually *adversely* affected these children's retest performance. Further support for this can be found in the observation that a larger proportion of failures to engage occured in the second session than in the first. Neither of these findings were statistically significant, however, and given the substantial number of changes occuring in the opposite direction, it would appear that something more than simply increased familiarity with the testing situation must be affecting children's willingness to perform to full competence on certain items.

What is significant about this pattern of variability is the fact that where performance on individual items was found to improve in the second session, these improvements were often confounded by the fact that they were achieved at the expense of items which children had shown themselves capable of passing in the first session. Consequently scores from neither session in fact reflected children's optimal or 'true' levels of ability. Interestingly findings from a second study by Horner (1988) revealed a pattern of results highly similar to those of the present investigation. From a more detailed analysis of the test-retest profiles obtained from his original group of 9 and 15 month olds, he was able to demonstrate a 50% probability that an infant's actual Developmental Quotient from a single examination will differ from his/her optimal score by more than one standard deviation of the BSID. Discussing these findings Horner asserts that

> "There is of course expectable error in any score produced by single examinations. No subject is ever expected to perform maximally during any given examination, and test behaviours of infants epitomise this fact. Elementary principles of sampling hold that a true score for an individual is the mean, and not the maximal score in a hypothetical distribution of that individual's scores under conditions unaffected by experience or fatigue. *Practically speaking however the clinician does not deal with a set of hypothetical scores but rather with a set of real scores (typically a single score) where experience, fatigue etc do play unavoidable roles in the derivation of these scores."*
>
> (my italics)

This latter point would seem especially pertinent in the case of mentally handicapped children. It has already been demonstrated that the overall level of variation in performance observed in DS children exceeded that found in Horner's non-handicapped subjects. This difference aside, the outcome of a below optimal assessment performance is likely to have far more serious implications for a handicapped child. If scores obtained from these children are not reflecting their optimal ability levels, it is very likely that any decisions made on the basis of these scores will be inappropriate.

The many failures to engage observed by DS children in the present study suggest that motivational factors play an important role in the variability of test scores in this population of children. This is not to suggest that similar variables do not affect the test performance of non-handicapped children. Given the levels of performance variability observed in Horner's study, it is likely that non-cognitive factors were also to some extent responsible for the discrepancy between optimal and obtained levels of performance in his subjects. The link between cognitive and motivational variables is increasingly being recognised to be an inextricable one, regardless of the ability level of the child. However the negative implications of these non-cognitive factors are likely to be more resounding for a child with mental handicap and not only in terms of decisions made on the basis of assessments. It is also possible that the adverse influence on performance of motivational variables may have wider-reaching effects on the development of cognition in the handicapped child (see below).

Suitability of the test items

The frequency with which children were seen to 'switch out' of BSID items is of significance on two other fronts. The first relates to the suitability of test items for use with older children with mental handicap who are older than the infants for whom these infants were first designed. As previously mentioned, data from psychometric tests is often used to support the theory that development in handicap is simply delayed in relation to normal development. Implicit in the use of such tests is the assumption that this 'slowness' applies not only to the rate at which children approach and reach developmental milestones, but to the pace of their lives in general. This point may best be illustrated with an example. On presentation with item 88 (picks up cup: secures cube; age placement 9 months; range 6-14 months), an 18 month old child with DS whose overall performance was well beyond this level flung both objects from the table and refused to cooperate on any subsequently presented items until, to pacify her, she was given the jointed doll. She was unable to follow the directions to put the doll on the chair, wipe its nose etc.

(item 126; age placement 17-18 months; range 14-26 months), nor would she point to parts of the doll's body (item 128; age placement 19.1 months; range 15-26 months), not, it appeared, because she was refusing to engage in these tasks, but simply because they were too advanced for a child at her level of cognitive functioning (optimal raw score 107; MA 13 months). Despite this, she clearly found the doll far more interesting than either the cup or the cube. This is hardly surprising given that she was 18 months of age. The test procedure however demanded presentation of item 88 to establish her basal level.

The BSID manual states that examiners should use their discretion when attempting to establish basal levels. It states that because infants can quickly outgrow specific reactions, failure to perform a specific task may indicate that the child has 'surpassed' that item and should therefore receive credit for it. This decision can generally be based on the child's response to higher level items in the same set. For example there are a number of tasks in the BSID which, like item 88, test the development of the object concept. To be confident that the child referred to above had surpassed item 88 it would be necessary to ensure that she had reached a higher stage in the development of this skill. Her performance on the other object concept tasks was ambiguous, however. She failed items 86 (uncovers toy) and 96 (unwraps cube), attending in both tasks to the tissue rather than to the object it concealed. Her performance on item 102 (uncovers blue box) seemed to be worthy of a credit, however; she was able to remove the lid of the box by toppling the box onto its side and she immediately reached for the object inside it in apparent indication that she understood the relationship between the two objects. To an examiner who had had no contact with this child prior to this testing session it would be almost impossible to establish from her performance, at what level she was capable of functioning. It would appear that she had 'outgrown' the objects used to test this skill but neither the test procedure nor the guidelines could accommodate this. Consequently, in the absence of sufficient information with which to determine whether or not she had surpassed item 88, it was not possible to credit her with a pass.

In constructing the BSID, test materials and tasks were selected on the basis of their proven suitability for use with non-handicapped infants in the target age ranges. Objects such as cups and cubes *do* hold the interest of the average 9 month old, both by nature of their relative novelty and their relevance in terms of the developing repertoire of cognitive and motor skills of children of this age. The age placements of the simpler items involving such objects precede those of the more complex tasks with the same objects by only a few months. The handicapped child, however, will inevitably have taken considerably longer to progress between these two developmental stages. In terms of day-to-day experience however their lives have not taken place in slow motion and there is no reason to assume that these objects should remain interesting for handicapped children for twice as long as they should for normally developing children.

Developmental implications of unreliable performance

Poor engagement in cognitive tasks, whatever its origins, must, if generalised to everyday learning situations, also have considerable implications for the progress of development. Although failures to engage were equally as prevalent on low level items, there were many instances of unreliability and avoidance towards the upper ranges of children's performance profiles. Given that in many cases children were also seen to experience genuine difficulty with other items at similar difficulty levels, failures by default and unreliable performance on these items are not as readily attributable to failure on the part of the items to capture the interest of these children. It may be that failures to engage in specific items are related to children's previous experience of failure on similar tasks.

It will be recalled that the cross-sectional study presented in the previous chapter indicated a possible link between items avoided by older children and those with which many younger childen had experienced difficulty. Although in itself this pattern of results does not directly indicate any longer term effects of early experience of failure, identification of similar links in the longitudinal performance of individual children would add considerable credibility to this hypothesis. Previous longitudinal research with 3-5 year olds and with infants with DS, using tasks designed to tap very early cognitive development, has indicated that the failures on these tasks by older subjects may not represent straightforward failures to acquire the requisite skills, but may in fact result from failures to consolidate these skills when they were first acquired in infancy (Wishart 1987,1988). Younger subjects were frequently observed to 'switch out' of tasks which they were known to have previously mastered, with performance typically falling away in the months following acquisition.

In both the Wishart studies and the present study, avoidance was observed on tasks at levels of difficulty above, or at the upper ends of children's overall ability level. Although perhaps less surprising than avoidance of easy tasks, this would suggest that DS children are creating further obstacles for themselves at the acquisition stages of learning. Unreliable or incomplete engagement at this early stage is likely to increase the probability of error. As indicated in the results from the errorless learning study (Chapter 2), experience of error is in turn likely to increase the probability of poor or below optimal performance. Reluctance to perform to full competence on BSID items may reflect the operation of the same effects. Difficulty and failure experienced at the acquisition stage of certain skills could result in a tendency to withold production of these skills even when they have been acquired. As the Wishart data suggests, this avoidance could then result in a failure to consolidate these skills.

If this spiral of initial failure - unreliable performance-failure to consolidate is characteristic of the cognitive development of children with DS, clearly DS developmental progress must be consistently undermined. Obviously a cross-sectional study such as the one presented in this chapter cannot provide any direct evidence on development or developmental processes. To explore in more detail the relationship between short-term unreliable performance demonstrated in assessment situations and the stability of skills in the longer term a longitudinal study of performance profiles is required. The following two chapters therefore report the findings of a study in which children were regularly tested and re-tested over closely spaced intervals for periods of up to 18 months.

CHAPTER 5

LONGITUDINAL PERFORMANCE ON THE BSID: A REASSESSMENT OF THEORIES OF DEVELOPMENT IN DOWN'S SYNDROME

A repeated finding throughout the two Bayley studies reported so far in this thesis has been that although on a quantitative level, the performance of DS children has been consistent with ability levels found in other DS studies, a very different picture emerges when the focus of analyses is shifted away from scores and other aspects of test performance are examined. Study 3 addressed the issues of reliability and validity in relation to the performance of young children with DS on the BSID. Item-item comparisons carried out on test and retest protocols of a group of DS children of varying ages revealed that performance in two closely spaced testing sessions could vary substantially. Although there were no significant differences found between scores attained in the two testing sessions, percentage agreement figures were found to be significantly lower than those found in similar studies with non-handicapped infants. An obvious implication of this finding was that neither testing session was able to provide an accurate estimate of children's 'true' levels of cognitive functioning.

Qualitative analyses indicated that low levels of reliability were in part attributable to the frequency with which children were observed to avoid the tasks presented to them in one or other session. In itself the prevalence of this type of 'cognitive avoidance' behaviour can be used to question the validity of this method of assessing the abilities of mentally handicapped children. Results were also discussed in terms of the suitability of items designed for and standardised on much younger non-handicapped children for use with children with mental handicap. It was argued that given the inevitable age differences between the two populations of children it should not be assumed that the same sets of items will be equally appropriate for eliciting demonstration of specific skills from both groups.

A great deal of research with DS and other mentally handicapped children has, however, been carried out under the assumption that psychometric tests of infant development, the BSID in particular, *can* provide accurate estimates of DS children's levels of cognitive functioning. Indeed various studies have been carried out comparing different psychometric tests of infant development in an effort to determine which are most appropriate for use with these particular populations (Ramsay and Piper 1980; Maisto and Germain 1986). A new test, the Battelle Development Inventory (1984), designed specifically for use with handicapped populations, has, moreover, been repeatedly investigated for reliability and validity against the BSID because of its "good psychometric properties" and because it is "considered one of the best instruments for assessing infant development, making it a good choice for use in (eg) a criterion-related study" (McLean et al 1987; see also Guidubaldi and Christie 1982; Guidubaldi and Perry 1984; Guidubaldi et al 1981; McLean et al 1987; Sexton et al 1988; Boyd et al 1989). Clearly, if neither statement can be substantiated with respect to the BSID when used with mentally handicapped children, nor can either be applied to any newly devised test which uses the BSID as a standard for comparison.

The purpose of the majority of the above studies has been to establish a reliable means of evaluating the effectiveness of early intervention programmes. Garwood (1982) has repeatedly questioned the suitability of using psychometric tests in this way unless the population under study has been included in the norming process. He advised against the assumption that, because these tests may be reliable and valid methods of determining the cognitive abilities of non-handicapped children, they will necessarily be equally so when used with populations of handicapped children.

In the present chapter these same assumptions about reliability and validity will be addressed in relation to the use of psychometric tests for devising theories and models of development in DS. Particular emphasis will be placed on studies of the rate of development in DS. A longitudinal study reporting both quantitative and qualitative analyses of BSID data will be presented and its results contrasted with the results of a number of longitudinal studies which have reported quantitative data only.

The organisational approach; hidden contradictions

In addition to its use in measuring 'typical' rate of cognitive development in DS, psychometric data has often been collected on cross-sectional samples for the purpose of making more generalised statements about this population. It is not a recent notion, however, that the exclusive use of such data for determining levels of cognitive functioning in DS is too limiting. In his critical review of studies of the educability of people with DS, Rynders suggested in 1978 that data from studies of non-cognitive aspects of development should be incorporated into a broader perspective on this issue.

Proponents of the organisational approach to development in DS claim to have "moved away from an approach emphasising the cognitive deficits of these children to a more broad-based developmental perspective" (Cicchetti 1987). Instead of repeated reports of the extent of cognitive retardation in this population, more recent investigations of children with DS have focussed on a number of different behavioural systems. These have included Piagetian stages and sequences of sensorimotor development (Cardoso-Martins and Mervis 1985; Cicchetti and Mans-Wagener 1987), organisation of the attachment system (Cicchetti and Serafica 1981), interrelationships among affective and cognitive development (Cicchetti and Sroufe 1976, 1978), self-other understanding (Beeghly et al 1986), and negativism (Spiker 1979). Findings from these studies, it is claimed, suggest that the development of young children with DS is "markedly coherent and lawfully organised" (Cicchetti 1987), and that it undergoes patterns and sequences which are highly similar to those observed in normal development.

The increase in research into such areas of development does represent a shift away from an emphasis on the cognitive deficit. Unfortunately, however, although the perspective on development in DS has widened, it has not widened sufficiently to loosen the hold of the paradigm of psychometric testing Because a large number of such studies still incorporate psychometric data - generally for purposes of matching handicapped and non-handicapped groups - the whole approach contains a major contradiction. Underlying the use of normative tests are a set of assumptions which are a legacy of the predominantly cognitive approach

- one such assumption being that psychometric tests are both a reliable and valid means of determining levels of cognitive functioning in both populations. Results from Chapter 4 strongly suggest that this assumption is inaccurate in relation to DS; children's patterns of passes and fails changed considerably over two closely spaced testing sessions. The prevalence of failures to engage uncovers a further assumption built into psychometric testing i.e. that children will perform to full competence. Clearly this was not the case with this sample of DS children. Many children were producing test scores well below their 'true' levels of cognitive functioning. Not only does this finding indicate that theories supported by MA matches of handicapped and non-handicapped groups must be flawed, it also exposes a fundamental circularity in the organisational approach. It is not possible to broaden the perspective on development in DS if, at the outset, this perspective is inherently restricted by the parameters of the psychometric approach. By definition, any 'broad' perspective should consider the influence of motivational and other non-cognitive factors on cognitive performance.

Rate of development in DS

Another major approach to the study of development in DS is based almost entirely on psychometric data. Over the last 50 years a number of studies have been carried out for the purpose of determining the rate of development in this population. Although this approach, unlike the organisational approach, does not itself claim to be broad, its overall objective seems to be to contribute to the development of early intervention programs and as such its implications are necessarily wide-reaching.

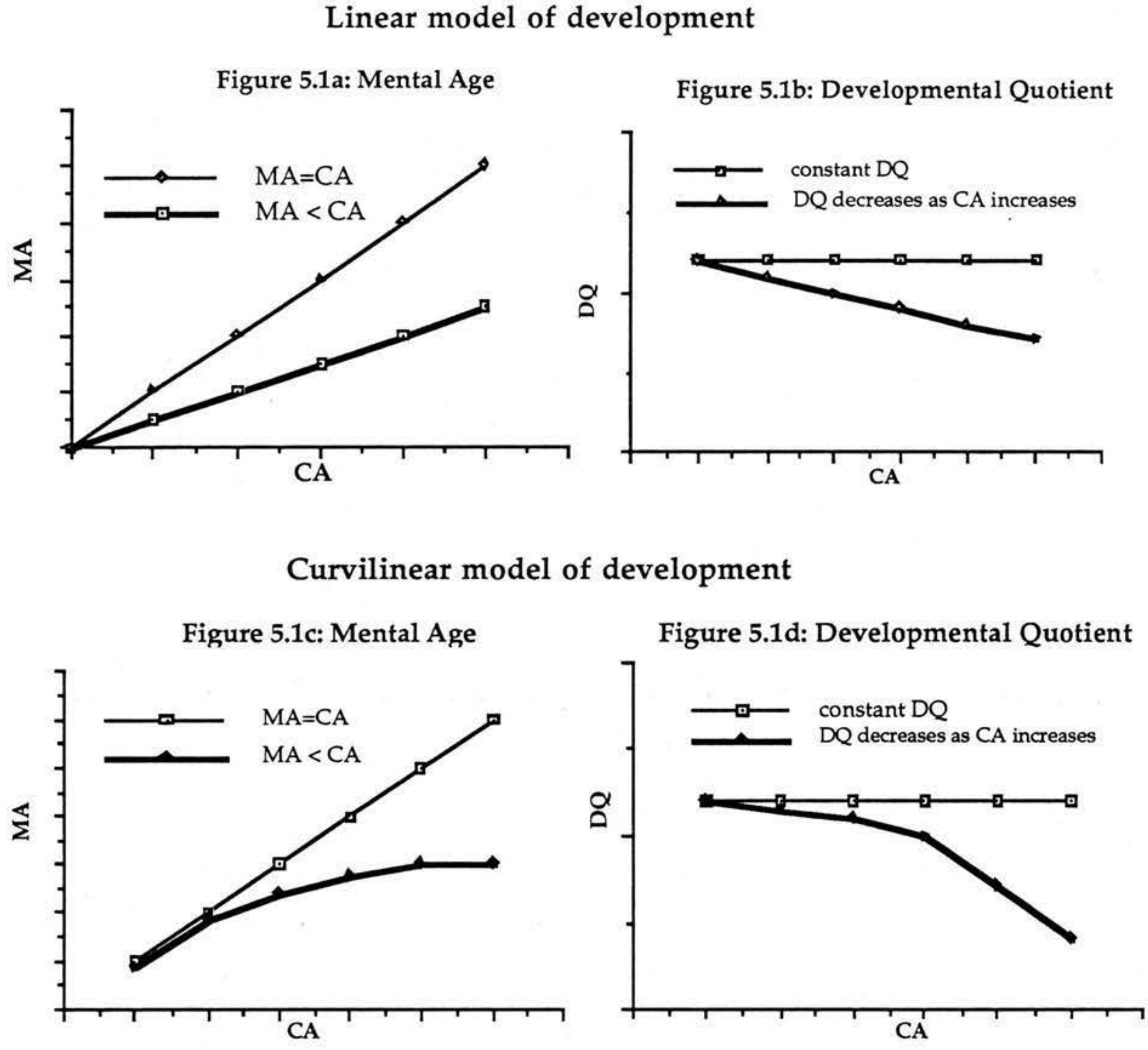
The majority of studies of the rate of development in DS have reported that the syndrome is associated with a decline in levels of intelligence with increasing age. With the exception of a study conducted by Kostrzewski (1974) almost all investigations of ability levels in this population have shown that, despite a steady increase in MA, IQ in DS declines with age. This decline has consistently been found in both cross-sectional and longitudinal studies conducted over the last 50 years (Gesell 1946 ; Oster 1953; Koch et al 1963; Loeffler and Smith 1964; Fischler et al 1964; Share et al 1961, 1964; Dicks-Mireaux 1966, 1972; de Coriat et al 1968;

Carr 1970, 1975, 1982; Eipper and Azen 1978,; Ramsay and Piper 1980; Hanson 1981; Schnell 1984; Piper et al 1986; Sharav and Schlomo 1988) and has emerged regardless of which tests are used (Ludlow and Allen 1979; Morgan 1979; Ramsay and Piper 1980).

Various controversies have arisen in the literature concerning the nature of the decline in DS IQ levels. One such controversy, particularly longstanding, relates to the exact nature of the decline: whether the rate of intelligence in DS simply declines relative to the normal rate (i.e. can be represented by a linear mathematical function) or whether it in fact decelerates (i.e. can be represented by a curvilinear mathematical function). Figure 5:1 a - d demonstrates these two different patterns of development. In a) development proceeds at a constant rate; with every increase in chronological age there is a proportional increase in mental age. Because CA is rising more rapidly than MA, DQ, which is calculated as a function of CA and MA, must fall. This decline is represented by the increasing divergence of the 2 curves in figure a) and can be seen more clearly when DQ is plotted against CA as in b). In Figures c) and d), however, it can be seen that the rate of development is not only slowing relative to the normal rate, but is also decelerating, i.e as CA increases the increments in MA made between each interval gradually reduce. This results in a downward curvilinear pattern of DQ as shown in d).

The majority of earlier studies lend support to a curvilinear model of development in DS. In 1979 Cunningham pointed out that despite a number of methological flaws and other shortcomings in investigations, a striking consistency among reports of intelligence in DS was evident with respect to the specific timing of the major period of this decline. Almost all studies reported a substantial decrease in IQ between the first and second year. The curvilinear function best fits this model of a delay plus early arrest. The only major opponent to this model was Dicks-Mireaux (1974) who found no evidence of a decline after 6 months and on the basis of these findings proposed a linear model of development in DS.

Figure 5.1: Linear and curvilinear models of development



It has been suggested that to some extent the debate over the nature of the intellectual decline in DS can be attributed to the contradiction between the decline in DQ and the simultaneous rise in MA. Carr (1975) has proposed, moreover, that the decline in DQ in the second year may in some cases be an artefact of the ratio of chronological to mental age. Cunningham (1979) however reported that even when the Dicks-Mireaux data were converted to MAs, the resulting curve still indicated a gradual slowing over time. This analysis of mental age was claimed to bypass artefactual possibilities associated with the use of DQs and "thus within the limits of the data available one must assume a relative decline in mental functioning" (Cunningham 1979). Several later studies, however, have found quotients to be relatively constant between 6 and 24 months. (Eipper and Azen 1978; Ramsay and Piper 1980; Schnell 1984)

Both linear and curvilinear models assume that the rate of development is being accurately monitored through the administration of single tests at each age level. Both assume that the scores attained at each age fully reflect childrens' overall levels of cognitive functioning and that any score increases made between ages accurately measure the extent of developmental progress made over that time. Both of these assumptions will be addressed in the longitudinal study to be reported below.

The linear model fits well with the 'slow' theory of development in handicap, with development proceeding steadily but at a slower pace than the normal rate. By contrast the curvilinear model implies that development actually decelerates. Both models have, however, been challenged by the discovery that DS children tend to reach 'plateaux' in development i.e. that for certain periods of time it appears that they do not make any developmental progress. The notion of plateaux was first proposed by Carr (1975) who noted periods when scores reflected little or no development having occurred between testing intervals. Rather than remaining constant, or decelerating slowly, she noted that some children's developmental levels showed large variations in both upward and downward directions. Carr suggested that the pattern of development in DS may not only be slower, but may also differ from that of normal children. Remaining longer on plateaux of achievements could explain the fluctuations in DQ levels. A reduction in DQ could be due to the fact that intervals between tests were not sufficiently long to span these extended plateaux. For example, if a child was tested towards the end of a plateau during which his/her developmental progress had been relatively stable for some time his/her DQ would seem to have fallen; if he/she were tested later having just made the step up from this plateau, his/her DQ would appear to have gone up.

Kirman (1974) tried to explain these developmental plateaux by suggesting that the anomalies in DS do not become manifest until the relevant stages of development are reached, and that these anomalies may be related to specific loci of damage. Cunningham (1979) subsequently found that errors in visually directed reaching tended to coincide with the commencement of the DQ decline. On the basis of this, he interpreted the decline and plateaux in development in DS in terms of specific inherent learning difficulties. In contrast, the more generalised notion of progressive deceleration would be represented in a developmental curve which would show relatively steady declines with little correlation to specific areas of development.

The discovery of developmental plateaux in DS may to some extent resolve the linear/curvilinear debate. It could be that studies reflecting the curvilinear (i.e progressive deterioration) pattern have contained a bias towards lower scores because of a proportionally larger number of 'plateauing' children being included in their samples. By contrast, a more constant rate of development would be reflected if larger numbers of children had recently 'stepped up' from plateaux, thereby maintaining the gap between mean chronological age and mean mental age at a more stable level.

A possible major flaw in the plateaux theory can be found in the fact, however, that its proponents used only quantitative data in their analyses. Plateaux were identified on the basis that scores did not appear to have increased over time. It will be recalled from Chapter 4 that many DS children were found to have attained very similar scores by very different routes. It may well be the case that plateaux can be explained in the same way - children may be producing similar scores over these apparent periods of no development, but these scores may have been achieved through very different patterns of performance. An obvious corollary of this could therefore be that plateaux are as much the result of failures to reproduce previous passes as from failure to acquire new skills.

The most recent international study of developmental rate in DS does not address the issue of plateaux. Instead this study focusses on the original linear/curvilinear debate which it may, to some extent, have resolved. This study involved BSID data collected from a sample of 229 children (age range 3 - 61 months) from 3 countries, West Germany, Australia and Canada (Rauh et al, 1990 in press). Using highly sophisticated statistical techniques attempts were made to fit this data base to both linear and curvilinear models of development. Although at a group level both models fitted the data tolerably well, random comparisons of individual children's BSID data resulted in the concession that 'there is insufficient theoretical and empirical evidence that sophisticated mathematical functions, linear or curvilinear represent best developmental growth'. From this it was concluded that it may be more appropriate to classify children according to the qualitative characteristics of their individual growth curves.

To a degree, in admitting that DS children cannot be considered to be homogeneous in terms of their rate of developmental progress, the approach of this latest study indicates a step in the right direction. A major shortcoming still however stems from the reliance on test scores for the purpose of determining individual growth curves. The following analysis of the results of a longitudinal study of DS childrens' BSID performance aims to demonstrate that until test performance is considered from a wider perspective than that permitted through simple comparisons of scores, few claims to an overall model of developmental progress in DS can be made. Rate of development will be considered through an item-item analysis of performance similar to that reported in Chapter 4. As in that study children will be tested twice at each age interval. This will serve two purposes: the first being to attempt to replicate the findings from the previous study in relation to 'optimal' scores exceeding scores attained in either of two closely spaced tests; the second to determine whether this method of score analysis has any effect on the resulting overall patterning of development in this group of DS children.

STUDY 4

METHOD

Subjects:

A total of 24 subjects participated in this longitudinal study, 16 females and 8 males. All subjects had already participated in the cross-sectional study reported in Chapter 3. Age on entry ranged between 3 months and 5 years.

Table 5:1 Longitudinal Study Sample

Age	3	6	9	12	15	18	21	24	27	30	33	36	42	48	54	60	Period*
1	2	2	2	2	2	2	2										18
2	2	2	2	2	2												12
3	2	2	2	2	2												12
4		2	2	2	2	2	2	2									18
5		1	2	2	2	2											12
6		2	2	2	2												9
7			1	2	2	2	2										12
8			1	2	2	2	2										12
9			1	2	2	2	2										12
10			1	1	2	2	2	2	2								18
11			1	2	1	2	2	2	2								18
12			1	2	2	2	2	2									15
13						1	1	2	2	2	2	2					18
14							2	2	2	-	2	2					15
15						1	1	2	2	2	2						15
16								1	1	1	2	2	2				18
17												2	2	2	2		18
18												2	2	2	2		18
19													2	2	2	2	18
20													2	2	2	2	18
21													2	2	2	2	18
22													1	2	2	2	18
23														2	2	2	12
24														2	2	2	12
A**	3	6	12	12	12	11	11	8	6	3	4	5	7	8	8	6	
B**	3	5	6	10	11	9	9	7	5	2	4	5	6	8	7	6	

1 = Tested once at this age level; 2 = Tested twice at this age level; 2 = Tested twice but score not included in group analyses (see text).
** Length of period followed in months.*
***A=> Number of subjects tested at least once at these age levels; B=> Number of subjects tested twice at these age levels.*

Procedure:

With the exception of one subject who was followed for only 9 months, all Ss were repeat-tested over periods of 12 - 18 months. The BSID was administered at 3 monthly intervals up to the age of 36 months and at 6 monthly intervals thereafter. In most instances Ss were tested twice at each age level (see below). Test and retest sessions were separated by 1 week for subjects aged up to 12 months and by two weeks beyond this age level. Table 5:1 shows the number of subjects tested at each age level and the number of sessions administered for each subject.

Due to funding and other difficulties encountered at the beginning of this study it was only possible to arrange single testing sessions with some children. In analyses where test and retest performances are directly compared therefore, scores and performance profiles are included only for subjects who had been tested twice at each age level.

Subjects 7 and 21 were tested twice at 12 and 54 months respectively but in both cases one of the two sessions had to be terminated before completion. These scores are not therefore included in the analyses of group scores. Performance profiles are included in the qualitative analyses, however.

RESULTS

Group Data

Table 5:2 shows the mean raw scores and corresponding MAs, MDIs and DQs attained in the first testing session at each age level. MDI levels below 50 were obtained from Naglieri's (1981) extrapolated indices. As these are only provided up to 30 months of age, scores attained beyond 30 months can be interpreted only as MAs or DQs.

Table 5:2: Session 1 mean Scores: Raw Scores, MDIS, MAS & DQ's

Age (months)	No. Subjects	Mean Raw Score	Mean MDI	Mean MA (mths)	Mean DQ
3	3	28	83	2-2½	75
6	6	50	66	4-4½	71
9	12	74	63	5-5½	58.3
12	12	82	46	7-8	62.5
15	12	89	40	9-10	63.3
18	11	97	36	10-11	58.3
21	11	106	39	13	62
24	8	109	39	14	58.3
27	6	117	45	16	59.2
30	3	134.3	60	20-21	71.6
33	4	131	-	19-20	59.0
36	5	130	-	19-20	54.1
42	7	135	-	20-21	48
48	8	144	-	24	50
54	7	148	-	25	46.2
60	6	156	-	30	50

Table 5:3 shows the differences in raw scores between each age level together with the corresponding changes in MA, MDI and DQ. It can be seen that the magnitude of raw score changes fluctuates between -3.3 and +24 raw score points. Because of the differing numbers of items presented at different mental age levels, however, fluctuations in the average number of items acquired between ages do not correspond proportionally to fluctuations in MA, MDI and DQ. MDI levels do nevertheless decline sharply twice within the first year and fall more gradually thereafter. These patterns are very similar to those obtained in previous studies of developmental rate in DS, suggesting a curvilinear model of development. Developmental rate is rapid in the first few months and seems to decelerate thereafter.

Table 5:3

Mean Raw Score and Mental Age Increases and Corresponding Changes in MD1 and DQ made between each Age Level

Age	Raw Score Increase	MA Increase	MDI Change	DQ Change
3-6m	22	2m	-17	-4.0
6-9m	24	1m	-3	-12.7
9-12m	8	2½m	-17	+4.2
12-15m	8	2m	-4	+1.2
15-18m	7	1m	-6	-5.0
18-21m	9	2m	+3	+3.7
21-24m	3	1m	-	-3.7
24-27m	8	2m	+6	+0.9
27-30m	17	5m	+15	+12.4
30-33m	-3.3	-1m		-12.6
33-36m	-1	-		-4.9
36-42m	6	1m		6.1
42-48m	10	3m		+2
48-54m	2	1m		-3.8
54-60m	8	5m		+3.8

Figures 5:2 - 5:4 present the group curves for MA, MDI and DQ when plotted against CA. MA rises steadily with increasing age. Excessively low scores attained at 18 months and excessively high scores at 30 months give the impression of a U-shaped MDI curve. As MDIs are not provided beyond 30 months it is not possible to demonstrate the overall decline in rate observed after this age level in MDI terms. This is more clearly represented in the DQ curve in Figure 5:3, however. With the exception of the minor interruption to the MA and DQ curves at 30 months, it would appear that a linear model, similar to that reported by Dicks-Mireaux (1974) most accurately describes developmental progression after the age of 12 months.

Figure 5:2: Session 1: Mean Mental Age Scores

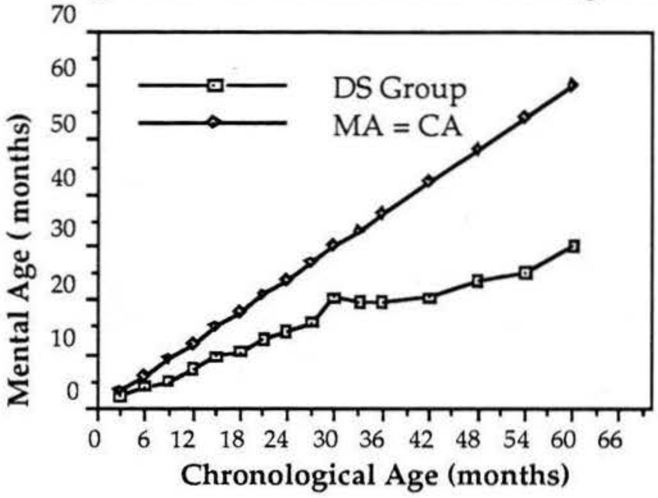


Figure 5:3: Session 1 Mean MDI Scores up to 30 Months

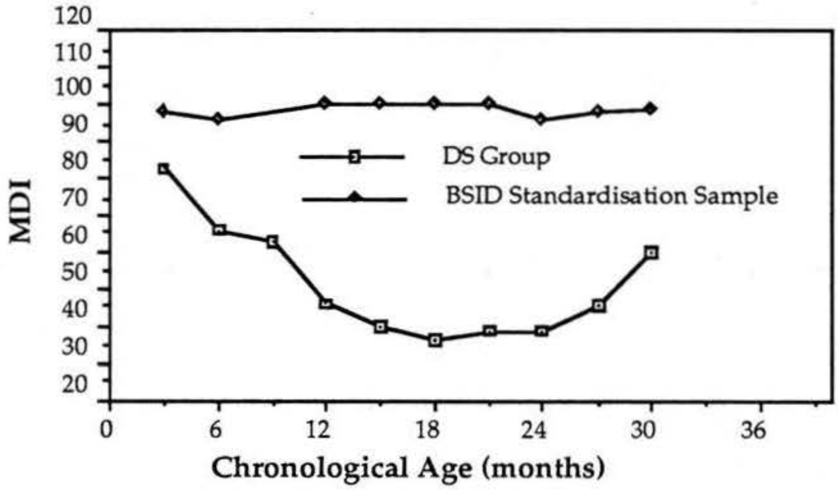
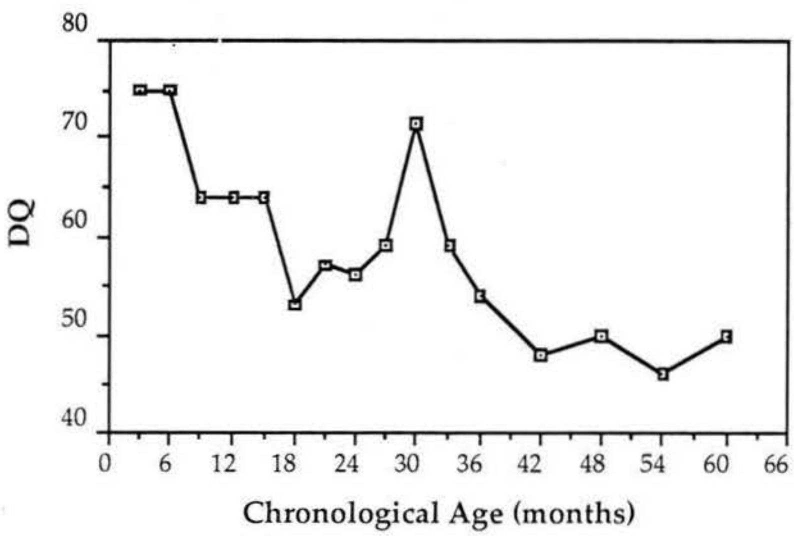


Figure 5:4: Session 1 Mean DQ Scores



Developmental curves of individual subjects

As noted in Chapter 3, an advantage of having a small sample is that mean figures do not obscure fluctuations in the overall pattern to the same extent that the means obtained from larger samples inevitably do. The predominantly linear nature of the group data was clearly at variance with the scoring patterns of many individual children and it was therefore decided to investigate the extent to which the group data might be masking individual score fluctuations. The growth curves representing scoring patterns of 6 children were examined. This group of children were randomly selected from the 12 who were followed for the longest priod, 18 months: their growth curves are presented in Figures 5:4 - 5:9. Scores obtained by Subjects 1 and 10 are presented as MDIs. MA curves are presented for Subjects 4, 18 and 19 because they did not attain sufficiently high raw scores to permit conversion into developmental indices. Subject 22's scores were converted into DQs to provide an example of a longitudinal pattern of scores expressed in this way.

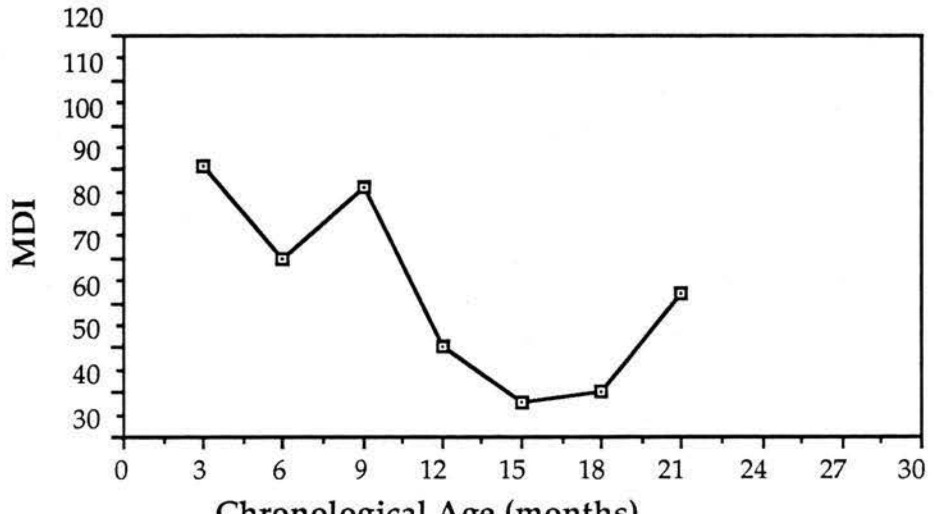
Subject 1: MDI (3 - 21 months) - Figure 5:5

Although the initial decline in this subject's MDI level between 3 and 6 months mirrors the group curve, the recovery of 16 MDI points at 9 months contrasts markedly with the group data at this age level. It can be seen from Table 5:4 that, rather than increasing, this subject's raw score actually decreased between 9 and 12 months and that this had a considerable effect on his MDI level which remained below 50 until 18 months. Between 12 and 18 months therefore this curve closely resembles the group curve. The sharp increase of 22 MDI points achieved at 21 months is not, however, reflected in the group curve; at this age level a much smaller increase in developmental rate is represented.

Table 5:4

Age (months)	3	6	9	12	15	18	21
Raw score	33	54	83	81	88	99	118
MDI	91	70	86	50	38	40	62

Figure 5:5: Subject 1 (3 -21 months): Session 1 MDI Scores



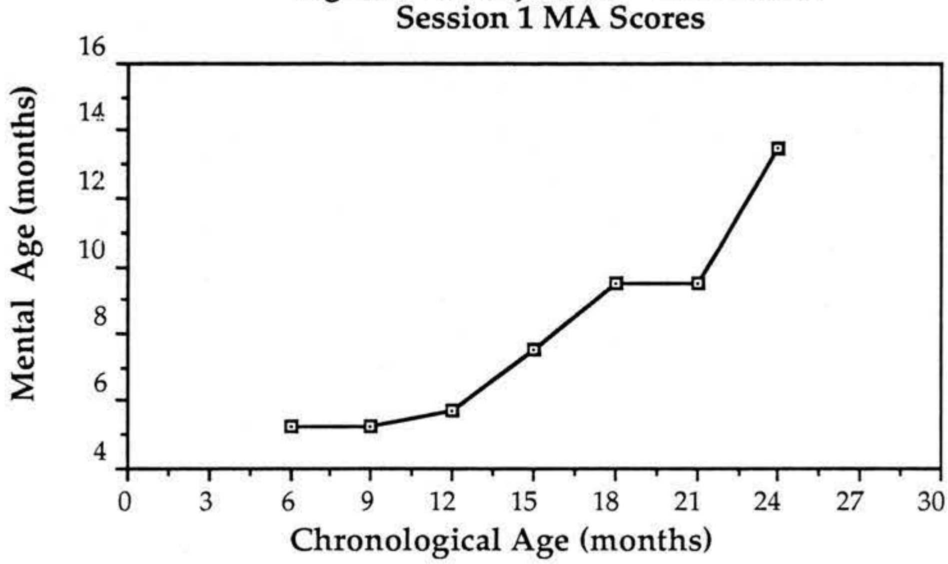
Subject 4: MA (6 - 24 months) - Figure 5:6

Two plateaux can be seen on this subject's MA curve, corresponding to a drop in raw score between 6 and 9 months and a very small raw score increase between 18 and 21 months. Neither are evident on the group curve. Between these plateaux, however, this subject's rate of development was proceeding close to the normal rate as indicated by the similarity of the gradient to that produced by plotting MA against CA. By comparison the DS curve from the group data changes direction slightly between 15 and 18 months, shifting away from the MA = CA curve and indicating an overall slowing in rate between these ages for the group as a whole.

Table 5:5

Age (months)	6	9	12	15	18	21	24
Raw score	62	61	72	80	89	92	108
MA (months)	5-5.5	5-5.5	5.5-6	7-8	9-10	9-10	13-14

Figure 5:6: Subject 4 (6 - 24 months)
Session 1 MA Scores



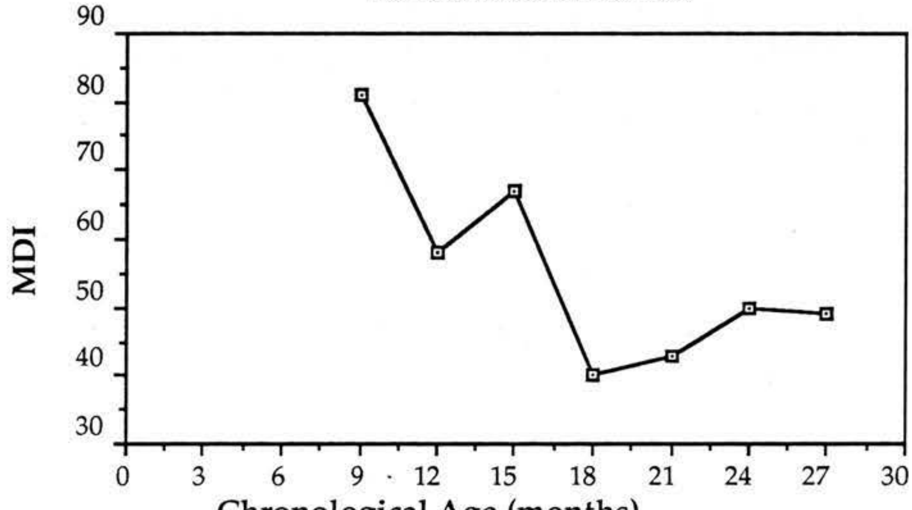
Subject 10: MDI (9 - 27 months) - Figure 5:7

Again, a somewhat erratic pattern of developmental rate indicating declines, inclines and a period of near stability of MDI level is shown by this subject. The sharp decline between 15 and 18 months corresponds to a loss of 2 raw score points between these ages. A similar, but less sharp decline is demonstrated on the group curve. This S's MDI rises by 10 points between 18 and 24 months however, in contrast to the much smaller increase of 2 points demonstrated on the group curve. In addition, whereas the group mean MDI level continues to rise beyond 24 months, a raw score increase of only 4 points corresponds to a small drop in MDI at 27 months for this particular subject.

Table 5:6

Age (months)	9	12	15	18	21	24	27
Raw score	81	87	101	99	108	116	120
MDI	81	58	67	40	43	50	49

Figure 5:7: Subject 10 (9 - 27 months)
Session 1 MDI Scores



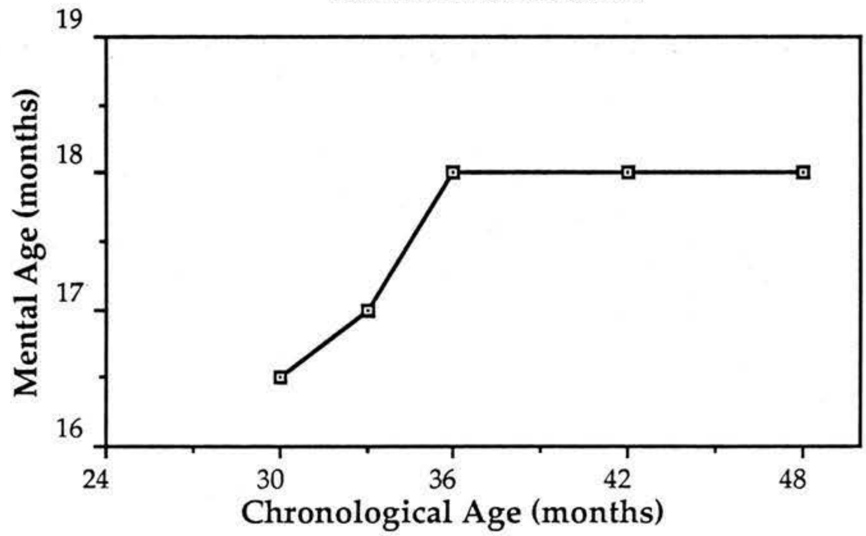
Subject 18: MA (30 - 48 months) - Figure 5:8

Clearly, this subject's BSID performance reflects little or no developmental progress having been made over the 18 month period during which she participated in this study. After an MA increase of 1-2 months between 30 and 36 months she remained on a score plateau for the next year. The group growth curve, by comparison, indicates a rise of over 4 MA months between 36 and 48 months.

Table 5:7

Age (months)	30	33	36	42	48
Raw score	117	120	124	125	126
MA (months)	16-17	17	18	18	18

Figure 5:8: Subject 18 (30 - 48 months)
Session 1 MA Scores



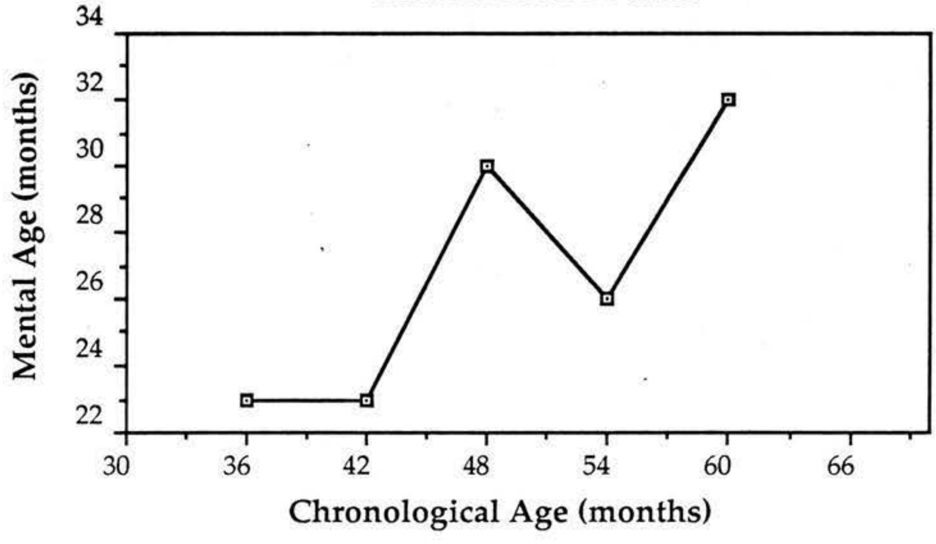
Subject 19: MA (36 - 60 months) - Figure 5:9

Between 36 and 42 months this subject gained only one raw score point, demonstrated by the plateauing effect seen between these ages. No such plateau is evident on the group curve. Also noteable is a loss of 6 raw score points between 48 and 54 months, resulting in a sharp decline of 4 MA months. The slight change in direction on the group MA curve between this same pair of ages may to some extent be attributable to this particular reduction in raw score. Only months later this subject had reached a ceiling on the BSID, however, and was scoring beyond an MA level of 30 months.

Table 5:8

Age (months)	36	42	48	54	60
Raw score	141	142	157	151	161
MA (months)	23	23	30	26	30+

Figure 5:9: Subject 19 (36 - 60 months)
Session 1 MDI Scores



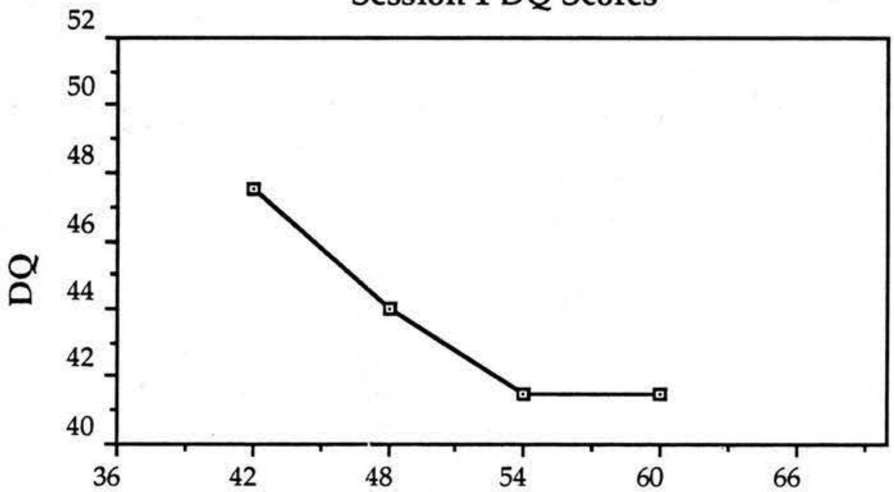
Subject 22: DQ (42 - 60 months) - Figure 5:10

A steady decline and eventual levelling out in DQ level is demonstrated in this subject's growth curve which quite closely resembles the group DQ curve. On the basis of scores from single sessions it would appear that this subject made steady but slow progress over this 18 month period.

Table 5:9

Age (months)	42	48	54	60
Raw score	133	138	141	148
DQ	48	44	42	42

Figure 5:10: Subject 22 (42 - 60 months)
Session 1 DQ Scores



Overall then, although the group figures do not themselves produce clearly linear developmental curves, at the same time they obscure a great deal of the variation which existed in individual growth curves. In general the group data produces a pattern in which MA rises and MDI and DQ fall with increasing age. This pattern is very much at variance with those produced by individual children's score profiles. As Rauh et al (in press) have suggested, it may be more appropriate to abandon the notion of an overall rate of development in DS and to focus more on individual patterns of growth.

Raw score reductions

In each of the individual cases reported above, MDIs and DQs were at times seen to decrease between various adjacent pairs of ages. This finding is consistent with the overall pattern found in this and other longitudinal studies of developmental rate in DS. A characteristic which has not as frequently been noted however is the frequency with which raw scores were actually seen to reduce over the 3 and 6 monthly intervals between testing sessions. Almost half this sample (11 Ss) were found to lose raw score points over these intervals on at least one occasion. Although many of these reductions amounted to only 1 or 2 raw score points, there were some considerable losses; these ranged between 5 and 13 points, translating into MDI losses of between 10 and 28 points. Given that even where raw scores have increased, gains have not been of a sufficient magnitude to sustain stable MDI levels, these further MDI reductions have a considerable effect on individual growth curves.

Cunningham (1979) also noted similar pass-to-fail patterns over successive 6 weekly assessments and remarked that this was more common for some items than others (this was also found in the present study and will be discussed in the following chapter). Children were only tested once at each age interval in the Cunningham study however. By retesting children at each age, it was possible to investigate whether these raw losses were present over two testing sessions, or whether they were simply the result of failures to perform to full competence in the first session.

It will be recalled from Chapter 4 that Ss were often seen to change their pass/fail patterns in both directions over closely-spaced intervals. Optimal scores calculated by combining all passes attained in both sessions were found to be higher than scores attained in either single session. Optimal scores were therefore again calculated and compared with the scores attained in the first of the two testing sessions.

As was the case in the study presented in Chapter 4, in almost every case where Ss were tested twice at the same age level in the present study, pass/fail patterns changed considerably (the resulting low item-item agreement figures will be discussed in Chapter 6). Table 5:10 compares the two sets of mean scores for subjects. Clearly, when all passes attained over the 2 sessions are credited this has a substantial effect on the overall pattern of raw scores. Mean optimal MDI and DQ levels are consequently substantially higher than those obtained in session 1; MAs increase by an average of 2 months.

Table 5:10 A Comparison of Session 1 Mean Scores versus Optimal Mean Scores

		Session 1				**Optimal**			
Age mths	No Subs.	Mean Raw Score	Mean MDI	Mean MA mths	Mean DQ	Mean Raw Score	Mean MDI	Mean MA mths	Mean DQ
3	3	28	83	2-2½	75	33.6	92	2½-3	91.6
6	5	53	69	4½	75	57.4	74	4½-5	79.1
9	6	71	55	5½-6	63.8	75.6	67	6-7	72.2
12	10	80	41	7-8	62.5	84	52	7-8	62.5
15	11	89	40	9-10	63.3	93.45	50	9-10	63.3
18	9	94	29	9-10	52.7	98.1	38	11-12	63.8
21	9	104	35	12	57.1	110.6	47	14	66.6
24	7	107	35	13-14	56.25	112	44	14-15	60.4
27	5	118	47	16	59.2	122.2	51	17-18	64.8
30	2	138	65	21-22	71.6	143	73	23	76.6
33	4	131	-	19-20	59.0	136	-	21-22	65.15
36	5	130	-	19-20	54.1	135.2	-	20-21	56.9
42	6	136	-	20-21	48	142.16	-	22-23	53.5
48	8	146	-	24	50	149	-	25-26	53.0
54	7	149	-	25-26	47.2	152	-	26-27	49
60	6	155		29-30	49.1	158	-	30+	50

Table 5:11 shows 10 cases in which Ss' raw scores in the first session were seen to drop between adjacent age levels and for which optimal scores attained over the same intervals were available. In the majority of cases it can be seen that optimal raw scores either increased or remained

the same over these periods. Three subjects' optimal scores however also dropped over the 3/6 monthly intervals, indicating that they failed items at a later age which they had passed 3 or 6 months previously. Moreover in the 2 cases where optimal scores did increase between adjacent age levels, these increases were very small: 1 and 4 raw score points respectively. This suggests that these children were not only underperforming in testing sessions, but that their overall developmental levels may have reached genuine plateaux. The small optimal score reductions made by subjects 8, 19 and 20 also indicate a possible plateauing of development in the intervals between 3 and 6 monthly tests, attributable not only to failures to pass new items but also to items dropping out.

Table 5:11 Raw Score Reductions over 3 and 6 Monthly Intervals compared with Optial Scores attained over the same terms

Subject No.	Age (Mths)	Raw Score/ Optimal Score	Age (Mths)	Raw Score/ Optimal Score
4	6	62/66	9	61/66
1	9	83/85	12	81/85
11	9	79/*	12	75/79
8	12	95/96	15	90/94 +
9	15	83/91	18	82/92
10	15	101/104	18	99/108
14	21	104/111	24	96/111
16	24	119/*	27	114/*
15	30	151/*	33	148/*
17	30	134/*	33	133/*
21	36	139/*	42	126/141
19	48	157/158	54	151/154 +
20	48	153/153	54	145/152 +

\+ *Despite use of optimal scores, raw scores still decreased between adjacent ages.*

* *Not tested twice at this age level.*

It will be recalled from Chapter 4 that Ss frequently attained similar scores via different pass patterns in closely-spaced testing sessions. It was therefore decided to investigate whether a similar phenomenon was in evidence for subjects whose scores had remained the same over 3 and 6 monthly intervals. An item-item comparison of the protocols of individual Ss whose optimal scores had either reduced, or reached plateaux between adjacent age levels was therefore carried out. Development was considered to have plateaued if scores increased by no more than 2 raw score points over the 3 or 6 monthly intervals.

Table 5:12 shows all cases of plateaux and score reductions. It can be seen that in 3 cases plateaux occurred between raw scores 80 - 105 (MA 7.5 - 12.5 months). This is the raw score level at which Cunningham (1979) found that plateaux were most likely to occur among his sample of subjects (who were aged between 18 weeks and 24 months). It can also be seen from Table 5:12 that in all cases where Ss' scores had dropped or had plateaued between adjacent pairs of ages, this plateauing effect was in part attributable to items having been dropped from childrens' protocols.

Table 5:12 Raw Score Plateaux Observed between Adjacent Age Levels

Subject No	Age(mths)	Raw Score	Age(mths)	Raw Score	No. Items Dropped
1	9	85	12	85	3
4	6	66	9	66	8
8	12	96	15	94	7
9	15	91	18	92	3
11	9	79*	12	79	4
14	21	111	24	105	6
15	18	121*	21	121*	3
15	30	156	33	149	8
16	24	119*	27	114*	6
16	33	132	36	131	5
18	36	128	42	127	4
18	42	127	48	128	4
19	48	158	54	154	5
20	48	153	54	152	4
24	54	156	60	158	2

* *Denotes cases in which subjects were only tested once at the relevant age levels. In these cases single session scores are compared. The remaining scores are optimal scores.*

For every subject an item-item comparison was carried out between optimal protocols obtained over every 3 and 6 monthly interval in order to determine the extent to which item loss might be influencing individual growth curves. In total there were 81 cases in which items which had been passed at a younger age were failed twice by the same children when tested at the next age level. In 35 cases the same items were again reliably failed a following 3 or 6 months later. In total 217 items were dropped at least once, with a further 51 items remaining absent from Ss' protocols at subsequent age levels. An average of 3.2 items were therefore lost in each case. This mean figure however obscures the magnitude of several individual cases of item loss in which as many as 8 items which had previously been passed were failed by the same child at a later age.

To exemplify the extent to which this phenomenon can affect the growth curves of individual children the scores of the 6 individual Ss whose first session growth curves were described above were amended to include all items which they had been observed to pass at least once during the period of time in which they had participated in this study.

Figures 5:11 - 5:15 demonstrate the effect on scoring patterns of these 6 Ss when:

- firstly, optimal scores are used at each age level, crediting all passes attained over the 2 closely-spaced intervals (optimal curve)
- secondly, all items passed at previous age intervals are added in to these optimal scores (optimal plus curve).

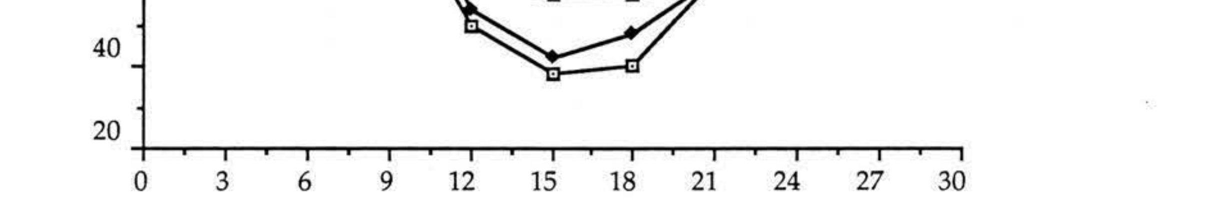
(In amending scores in this way it is not intended to imply that the resulting scores should be taken as 'true' indicators of levels of cognitive functioning. Clearly, if skills are being lost from children's repertoires it would serve no useful purpose to overlook this. The objective of this particular investigation is to explore the nature of the decline which seems to characterise the rate of development in DS. It may be that the loss of skills is actually contributing to this decline).

Subject 1: MDI (3 -21 months) - Figure 5:11

A total of 10 items were dropped from this S's repertoire of passes - 3 between 9 and 12 months, (which were also failed at 15 months) 4 between 12 and 15 months (2 of which he failed again at 18 months), and a further 3 between 15 and 18 months.

The most dramatic difference between the session 1 curve and the optimal-plus curve can be seen between 12 and 18 months. At 12 months, by including all items previously passed, this S's raw score is enhanced by 7 points, translating into a 17 point difference in MDI level. At 15 and 18 months a raw score difference of 9 points at both ages corresponds to a difference of 20 and 18 MDI points respectively. When this S's scores are amended in this way it can be seen that the magnitude of the MDI decline between 9 and 15 months, although still substantial at 35 points, is 9 MDI points less than the drop from 82 to 38 seen in the optimal curve. Moreover the optimal-plus curve would indicate a more constant rate of development between 9 and 21 months, rather than the steep incline seen in the session 1 curve.

Figure 5:11: Subject 1 (3 - 21 months)
Session 1 MDI Scores Versus Amended MDI Scores

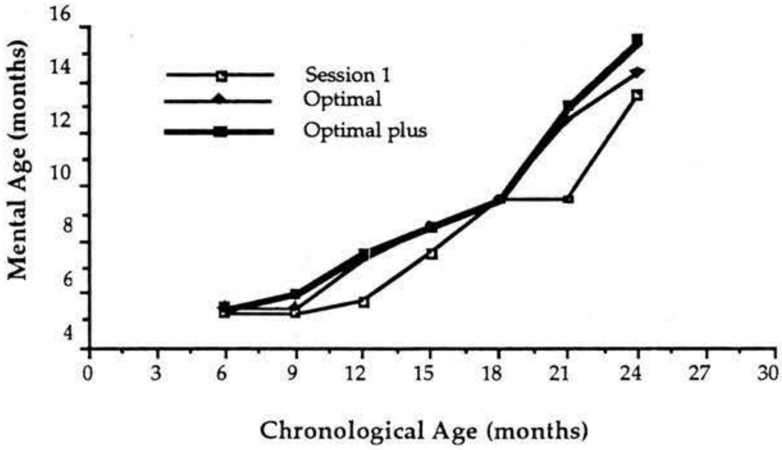


Subject 4: MA (6 - 24 months) - Figure 5:12

Clearly, with this subject, the major effect of amending scores is to almost abolish the 2 plateaux previously existing in her developmental curve between 6 and 9 months and between 18 and 21 months. Between the two earlier ages this was achieved by including items which had dropped out. Session 1 scores indicated a loss of 1 raw score point between 6 and 9 months (6 months: raw score 62; 9 months: raw score 61). Optimal scores still, however, indicated a plateau at a slightly higher raw score of 66. These simple score comparisons however obscure the fact that although 7 new items were passed between 6 and 9 months, a total of 8 items were also dropped over the same period. When these 8 items are included in the 9 month score (optimal-plus curve), it can be seen that rather than remaining constant, this subject's MA does increase slightly.

Between 18 and 21 months the plateau observed between session 1 scores was abolished by plotting optimal scores. Comparison of session 1 scores attained at these 2 age levels revealed a raw score increase of only 3 points. This small increase made no difference to this S's MA which remained at 9.5 months. By comparison, the difference in optimal scores between 18 and 21 months reached 14 raw score points, pushing the MA level up to 12.5 months at the later age, a rise of 3.5 months. By including the additional 3 raw score points corresponding to the 3 items which this subject had dropped between these ages it can be seen from the optimal-plus curve that her MA level is slightly further increased.

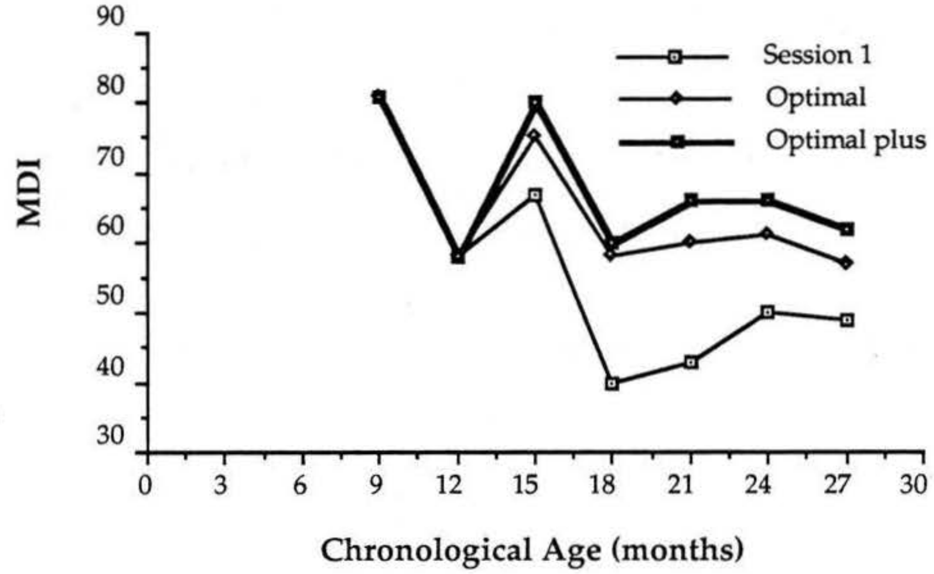
Figure 5:12: Subject 4 (6 - 24 months) Session 1 MA Scores Versus Amended MA Scores



Subject 10: MDI (9 - 27 months) - Figure 5:13

Because this subject was only tested once at 9 and 12 months it is not possible to present optimal scores for either of these ages. However, the effect of including all passes attained in either of the two closely-spaced sessions can clearly be seen at 15 months. This effect is most pronounced at 18 months where the session 1/optimal score difference amounted to 9 raw score points - or 18 MDI points. Comparison of the session 1 curve and the optimal-plus curve shows that the largest decline - between 9 and 18 months - is reduced by 20 MDI points if all previous passes are included in the score at 18 months.

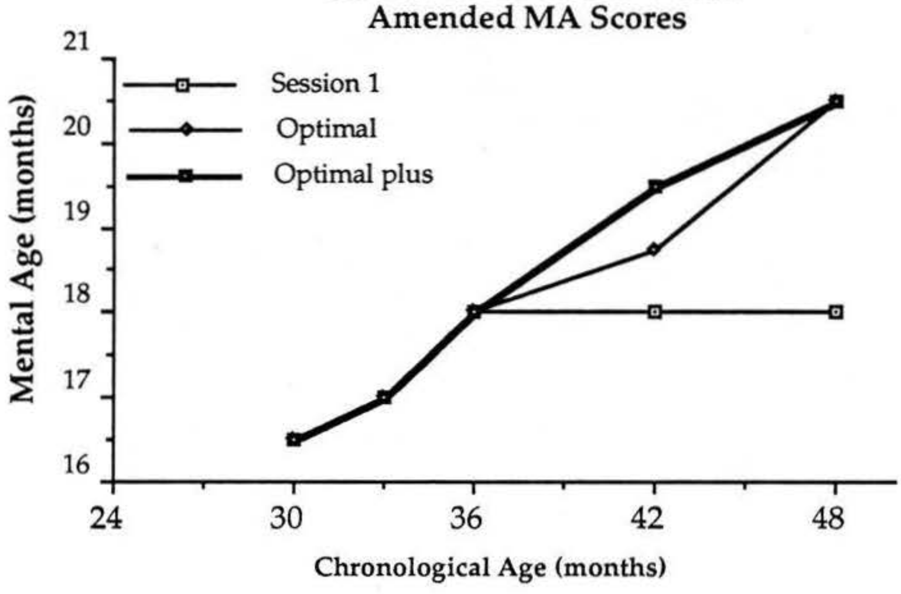
Figure 5:13: Subject 10 (9 - 27 months)
Session 1 MDI Scores Versus
Amended MDI Scores



Subject 18: MA (30 - 48 months) - Figure 5:14

As in Figure 5:11 it can be seen that the plateau in this S's session 1 curve, between 36 and 54 months is abolished when optimal scores are plotted. Rather than making no developmental progress between these ages, as would appear to be the case from session 1 scores, this subject was in fact developing at a rate almost consistent with that seen between 30 and 36 months. When all previous passes are credited as in the optimal-plus curve, development would appear to be progressing at a near constant rate.

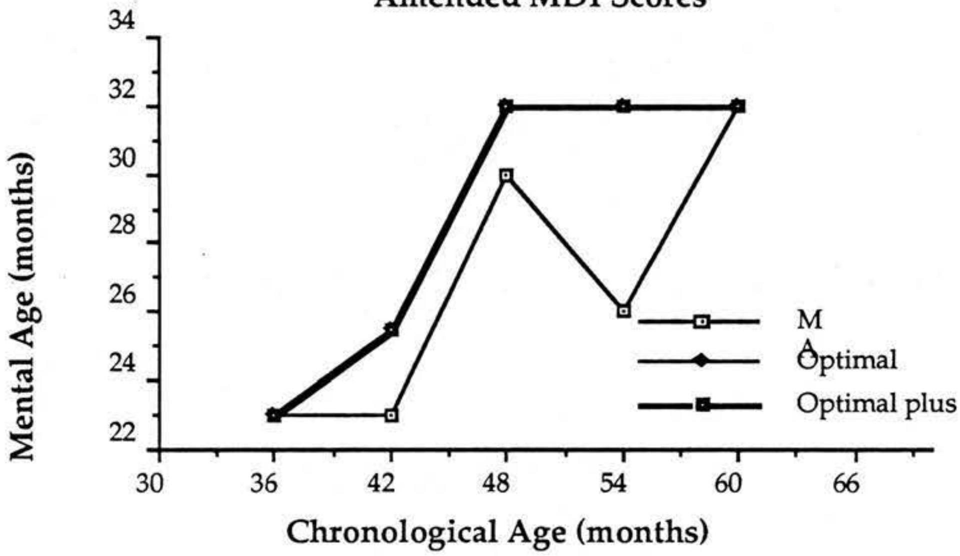
Figure 5:14: Subject 18 (30 - 48 months) Session 1 MA Scores Versus Amended MA Scores



Subject 19: MA (36 - 60 months) - Figure 5:15

Again the effect of plotting optimal scores is to abolish the plateau between 36 and 42 months seen in this S's session 1 curve. She in fact reached a ceiling on the BSID at 48 months, 12 months earlier than appears to be the case in the session 1 curve. Although this subject did drop several items from her performance repertoire at 48, 54 and 60 months, their addition to optimal scores made little difference to her MA pattern.

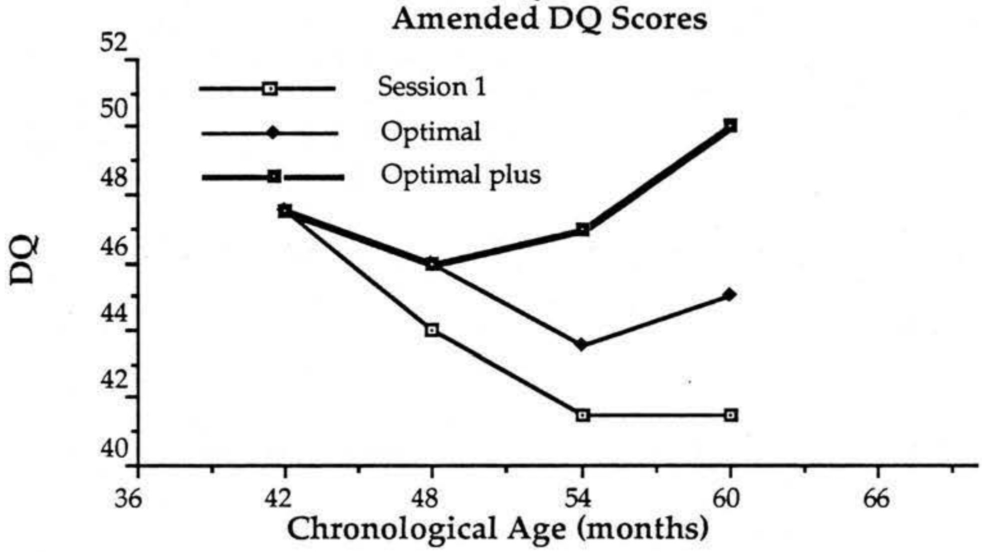
Figure 5:15: Subject 19 (36 - 60 months)
Session 1 MA Scores Versus
Amended MDI Scores



Subject 22: DQ (42 - 60 months) - Figure 5:16

In this graph both session 1 and optimal curves demonstrate a decline in DQ. This decline is not present, however, when all items previously passed are added to this S's scores as shown in curve c/. He dropped 5 raw score points between 48 and 54 months, corresponding to an DQ difference of 4.5 points. At 60 months the difference in DQ made by a gain of only 2 raw score points amounts to over 4 points. Although these differences are extremely small they do have the effect of producing a rising DQ curve, rather than one which declines steadily.

Figure 5:16: Subject 22 (42 - 60 months)
Session 1 DQ Scores Versus
Amended DQ Scores



Transformation of group curves: Although sizeable effects are produced on individual growth curves by including all items previously passed, this effect is substantially diminished when the overall group curves are amended in the same way. Figures 5:17 - 5:19 show the difference made to group MA, MDI and DQ curves when mean optimal scores are plotted and when all items which Ss had dropped between age levels are included in scores. In each case, although, inevitably, scores are slightly higher at each age level, the general shape of the curves remains largely unchanged. MA rises and DQ and MDI decline at almost equivalent rates in curves a/ and c/ in all 3 graphs. This contradictory pattern seen between individual and group results substantiates the claim that it may simply be inappropriate to describe developmental rate in DS in terms of a simple mathematical function.

Figure 5:17: Session 1 Mean MA Scores Versus Amended Mean MA Scores

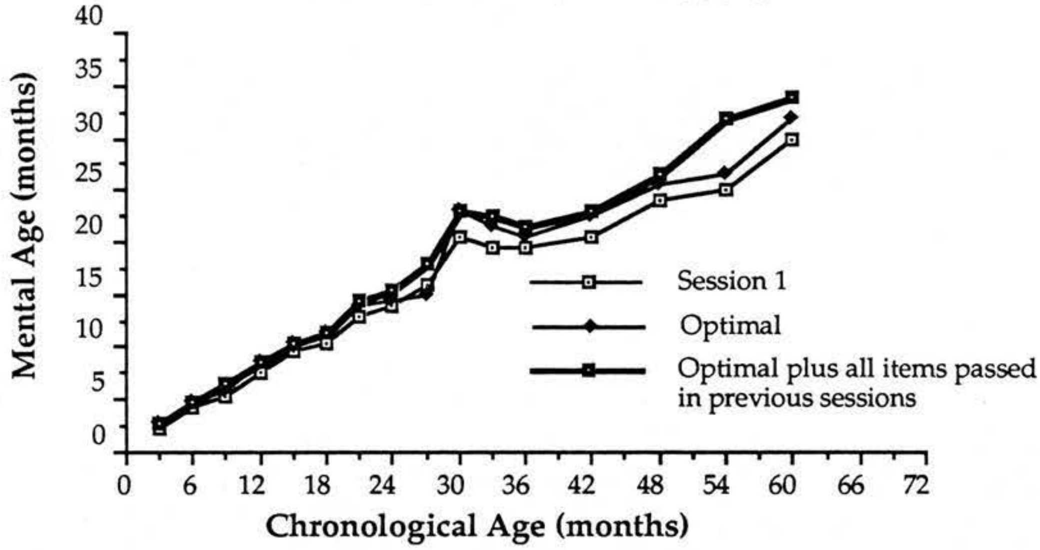


Figure 5:18: Session 1 Mean MDI Scores Versus Amended Mean MDI Scores

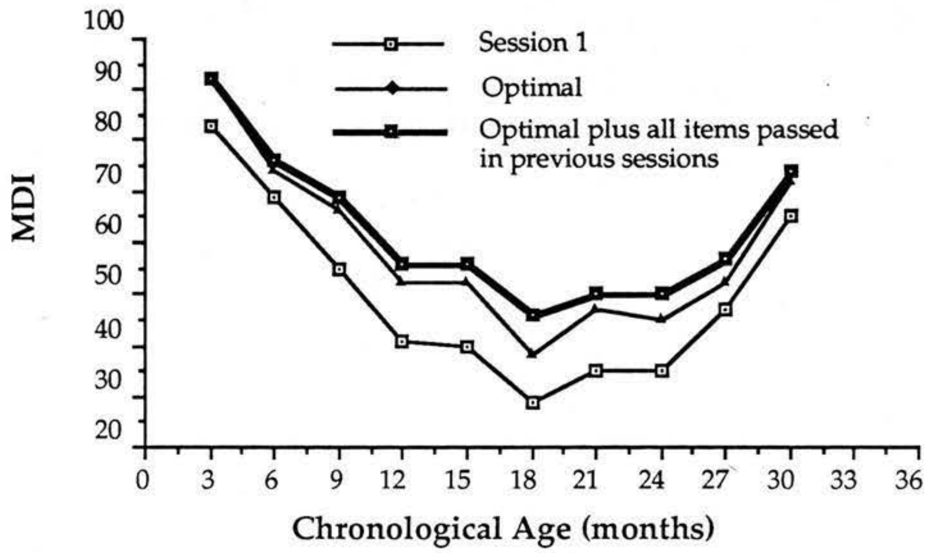
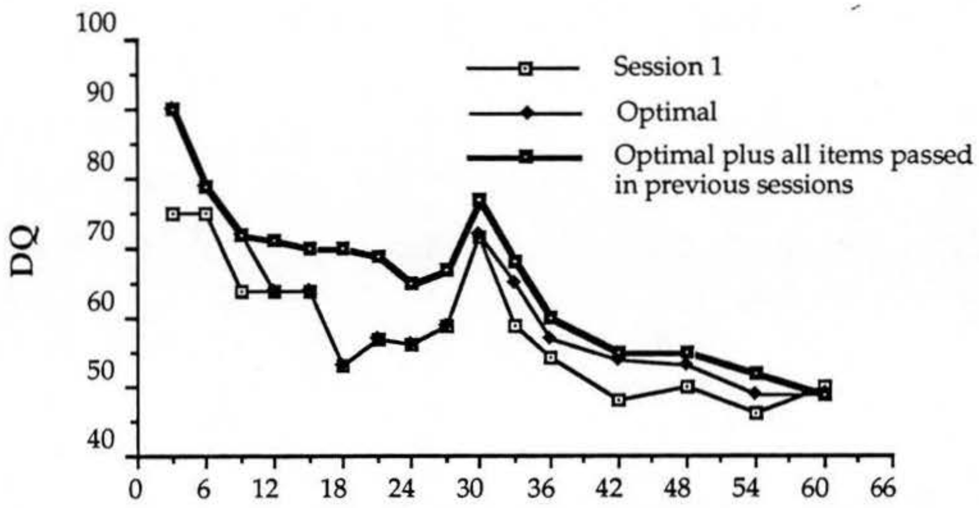


Figure 5:19: Session 1 Mean DQ Scores Versus Amended Mean DQ Scores



DISCUSSION

This study reported the findings of a longitudinal study of the performance of a group of children with DS on the BSID. Children were tested twice at each age level and in depth and individual as well as group quantitative analyses were carried out on performance protocols.

The overall pattern of developmental rate generated from only the first of the two testing sessions at each age level was generally consistent with that reported in previous longitudinal studies of DS infants and children. At 3 months scores were just over 1 sd below the standardisation mean but within the first year average performance fell rapidly against the norm. After this period of rapid decline, developmental rate levelled out, gradually falling behind the normal rate but proceeding at a largely constant rate up to 60 months of age. A linear model therefore most closely describes the developmental progress of this group.

Comparisons of individual Ss' growth curves of first session performance, however, revealed many idiosyncratic periods of change which were obscured in the group curve. The developmental rate of these individual children did not proceed at a constant rate but showed many fluctuations in rate between adjacent pairs of ages; in several cases little development appeared to have occurred within the 3 and 6 monthly intervals between testings. The prevalence of these variations in rate between ages suggests that any attempt to describe development in DS in terms of a single mathematical function is to some extent fruitless. This conclusion fits with that of the authors of the most recent large-scale study of DS BSID performance, the results of which are only now being published (Rauh et al in press).

One pattern relating less to age than to the magnitude of score increases between ages was in evidence around raw score levels 80 - 105. Rate of development did appear to decelerate as children reached this particular mental age level (7.5 - 12.5 months). As already mentioned, Cunningham (1979) reported a strong plateauing effect around this raw score level. In that study Ss were tested at 6 weekly intervals; the likelihood of detecting such periods of little or no developmental progress was therefore greater than in the present study.. Because of the wide range of chronological ages (9 - 24 months) at which children in this sample were performing at this mental age level, this pattern is not clear when MA is plotted against CA. It does, however, have the effect of lowering group MDIs between 9 and 24 months. The suggestion was made in the introduction that the linear/curvilinear debate may be in part attributable to the differing numbers of 'plateauing' Ss being included in samples. By excluding the age levels at which plateaux were most prevalent in this sample it is possible to demonstrate this effect. One of the most rapid periods of decline occurs at 12 months when almost half the sample (10) have reached the MA level covered by items 80 -105. Although, beyond this point, the decline is more gradual, a definite slowing in rate is evident as children approach this stage in the test. If testing intervals had been sufficiently large to by-pass this particular period of slow progress, there would have been little evidence to suggest that developmental rate had decelerated over this period.

Any description of development in DS which aims to be accurate should not, however, by-pass what would appear to be a particularly distinctive stage in the development of this population of children. Clearly, if developmental progress is genuinely slowing at this point it would make little sense to overlook this anomaly. Over an MA range of 7.5 to 12.5 months, the developmental rate of many childen with DS does not appear to be progressing at a rate consistent with that seen at both earlier and later stages in development, whether or not a linear model can conveniently be fitted to the group data as a whole.

The data of Rauh et al demonstrates just how this period of slowing developmental rate can be obscured by grouping together a large data set. (see Figure 5:20) Mean MA scores for their entire international sample, consisting of 229 DS children, produce a clearly linear progression, one which closely resembles that produced in a study of the BSID performance of 83 DS children carried out in the U.S.A. (Pueschel 1984). As well as plotting mean scores, Rauh et al also plotted lowest and highest scoring Ss separately. Curves representing these scores reveal a relative slowing in rate as MA approaches 12.5 months. Indeed it would appear that the mean MA level of lowest scoring Ss in this sample actually declined from 15 months to just over 10 months between 30 and 35 months. Both these changes in direction are however masked in the mean MA curve because, as with the sample used in the current study, MA ranges at each CA interval were wide. The narrower MA ranges which inevitably emerge from separate analyses of maximum and minimum scores, by contrast, result in a clearer picture of the CA/MA relationship. Moreover, where individual growth curves are presented for children in the Rauh et al sample, it is possible to see a number of plateaux as Ss reach MAs between 7.5 and 12.5 months. (Figures 5:21 - 5:23)

It has been suggested that the prevalence of these apparent plateaux in development may simply be artefactual, a function of the way in which the BSID has been constructed. The 'step size' between individual items (as measured by mean rate of acquisition in the standardisation sample) necessarily relates to rate of development. Over periods when normal development proceeds at a rapid rate therefore, the step size between items increases proportionally. Plateaux may simply exaggerate the difference in rate between normal development and the slower progress of development in DS over such periods.

Figure 5:20. Rauh et al (1990)

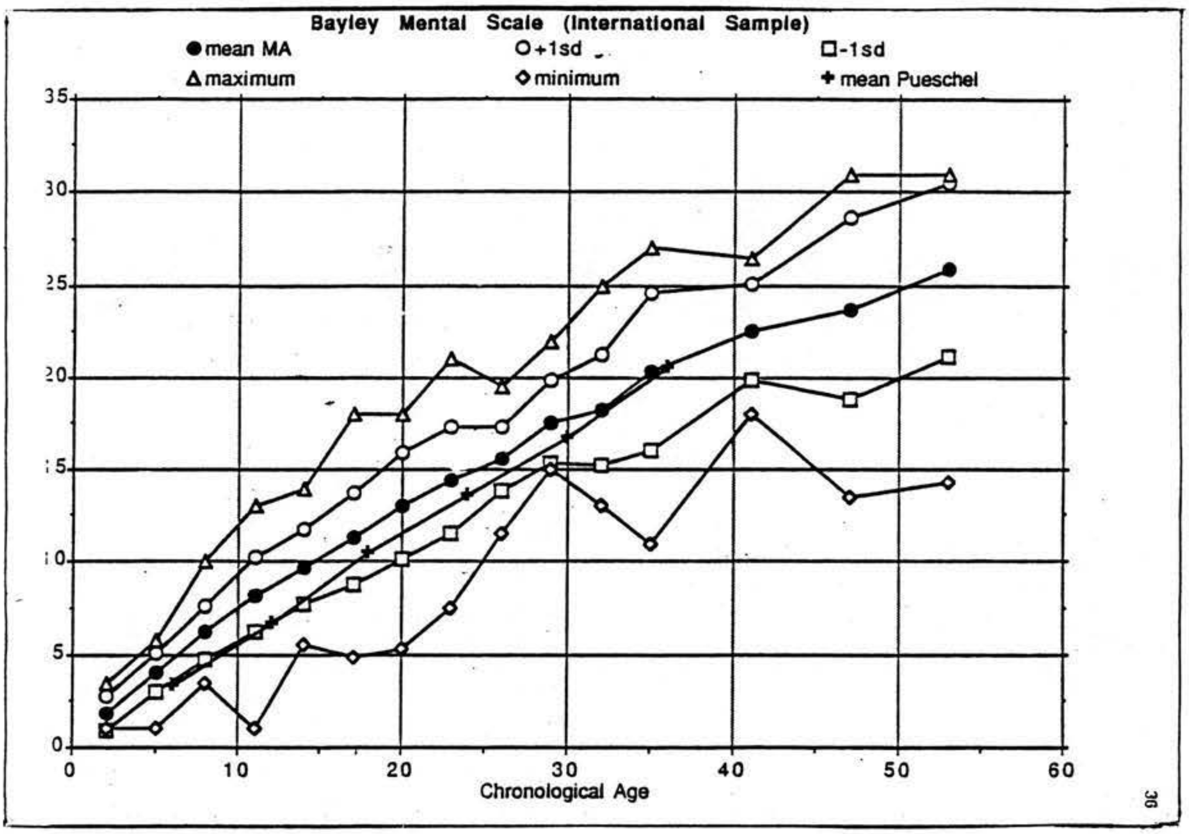


Figure 5:21. Rauh et al (1990)

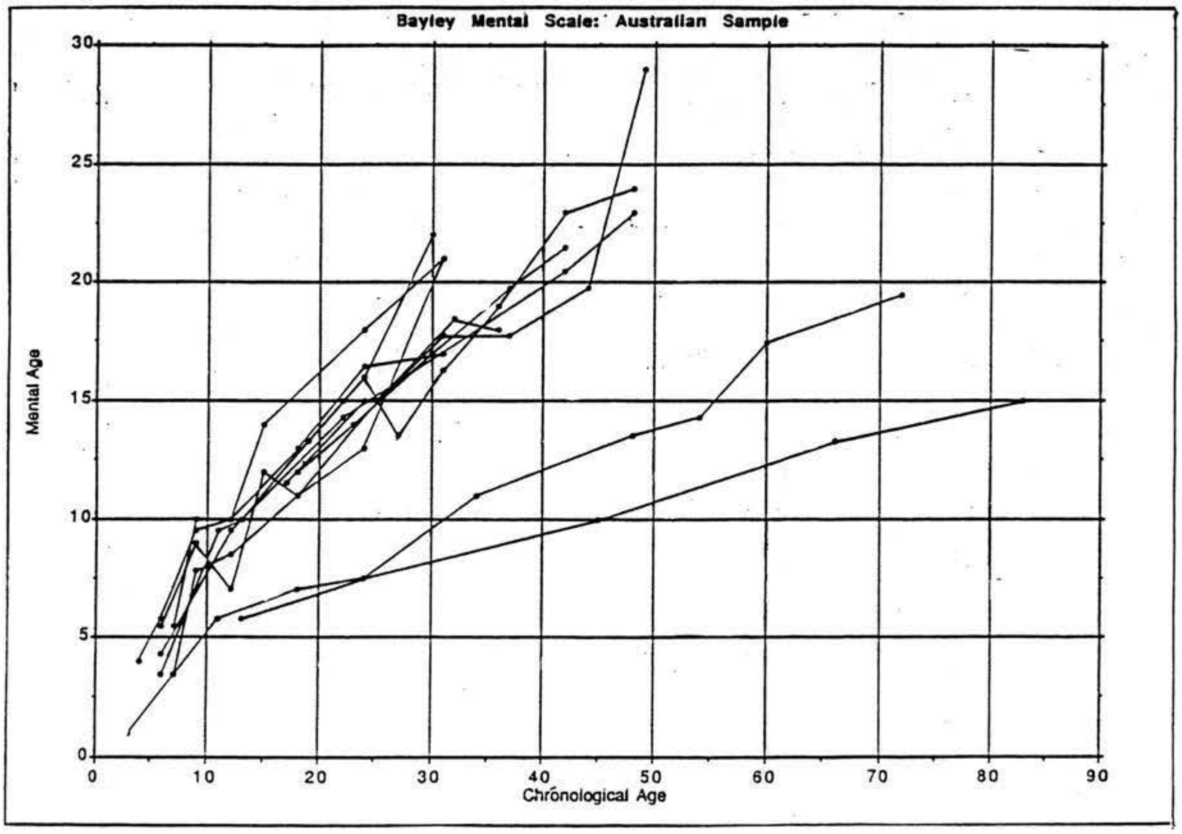


Figure 5:22. Rauh et al (1990)

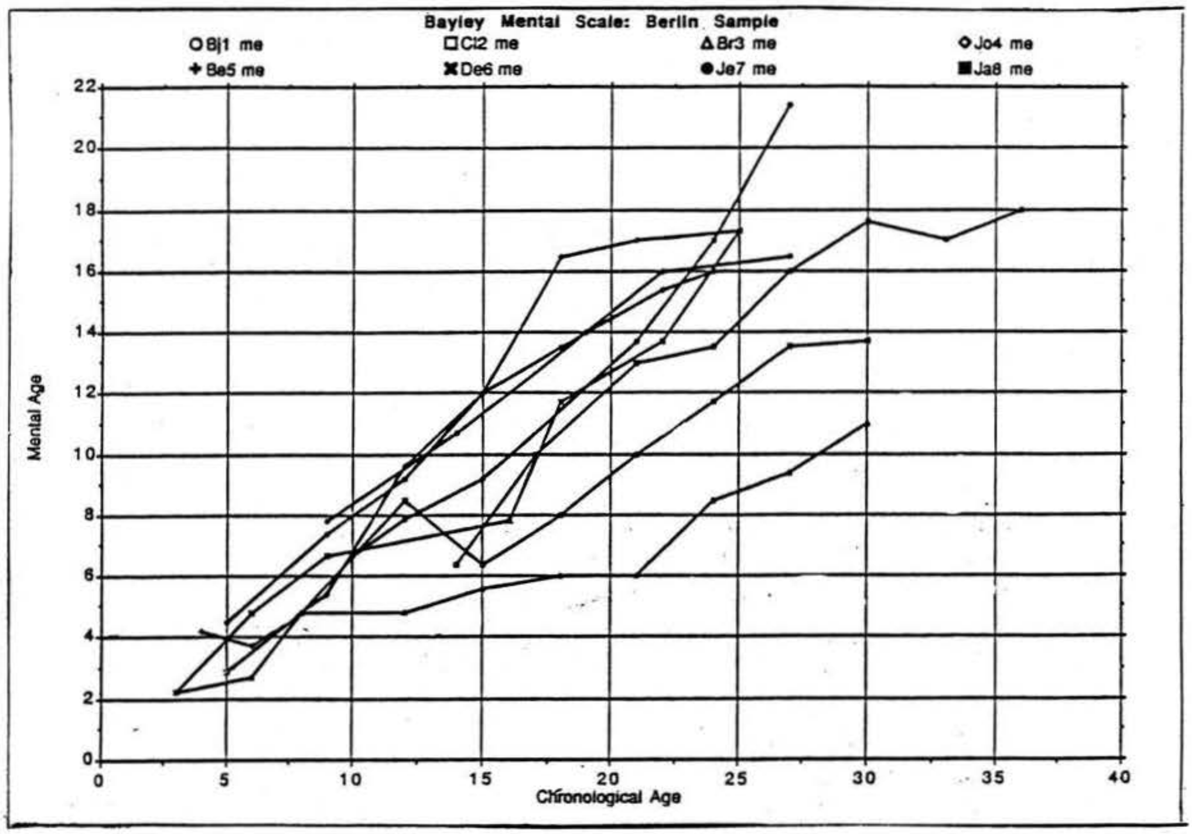
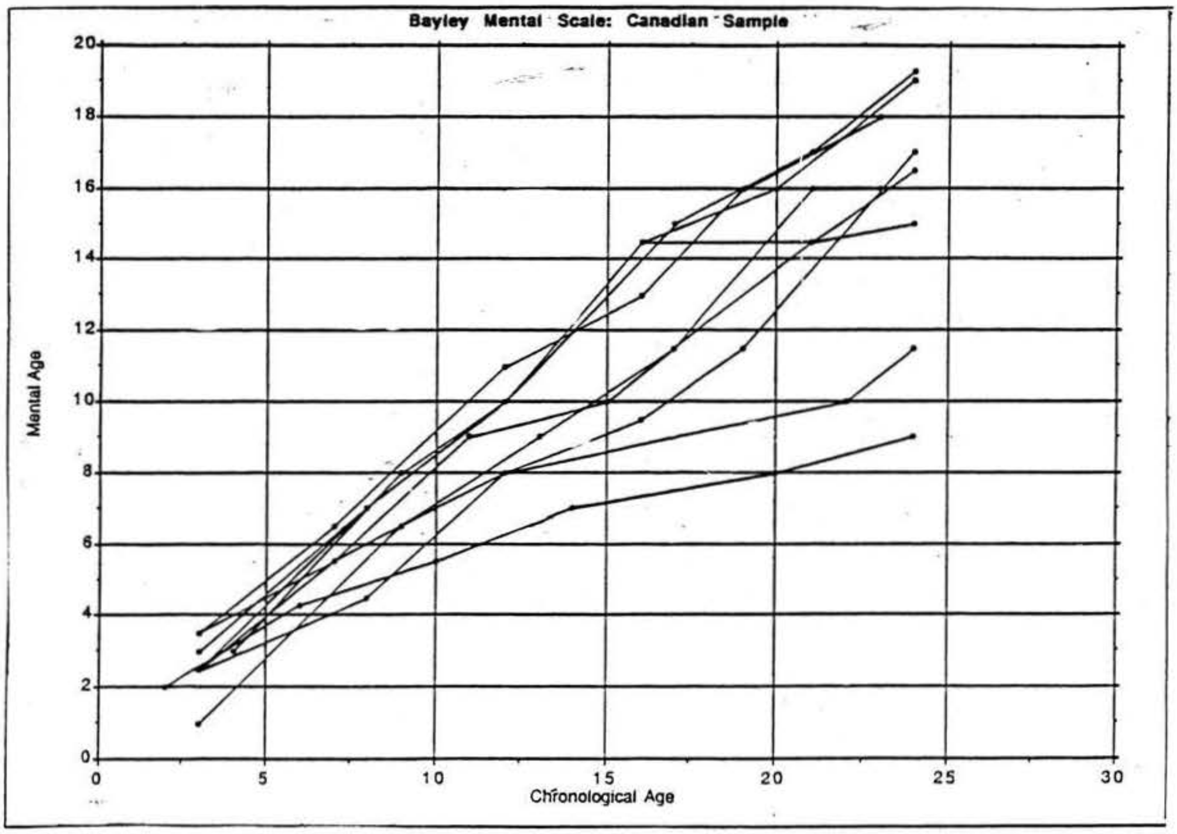


Figure 5:23. Rauh et al (1990)



Cunningham rejected this possible explanation for two reasons: firstly, he noted a prevalence of plateaux on the BSID motor scale at a stage in the test when relative step size was constant; secondly, some of the most frequently failed items falling between items 80 - 105 do not have larger than average step sizes. As an alternative explanation Cunningham supported Kirman's (1974) hypothesis that developmental plateaux reflect specific deficiencies in DS which become manifest only when the relevant stages in development are reached. Data from the present study, however suggests that a more complex explanation is required for these periods of little or no developmental progress - one relating as much to motivational factors as to deficiencies directly attributable to the handicapping condition.

By comparing the percentage failure rates of individual items, Cunningham noted two areas in which specific difficulties seemed to appear in early development and suggested that these may indeed predict plateaux at later stages. One such area of difficulty related to the coordination of vision and fine motor movements in early reaching, the other to sustained attention and the beginning of more purposive behaviour. At this age level there are a number of items (e.g item 68: exploitive string play) which can be passed despite low levels of activity on the part of the child. Normally developing infants usually demonstrate these behaviours with a high level of excitement and activity, however; parameters such as speed of response and level of activity are assumed and not directly tested. DS children, by comparison, often demonstrate particularly low levels of activity when presented with such items, particularly for the first time, although their responses do fall within the criteria for passing. When tested when slightly older, however, the same infant may produce a more 'normal' response. Cunningham suggests that by just passing items at criteria levels, DS infants may create a plateau situation. Not only must the infant bring these specific behaviours 'up to strength', but to raise his/her score, new items must be attained. One implication of this is that to test such behaviours adequately in DS infants, items should be more precisely defined in relation to their components. This would lead to a series of sub-items, in turn reducing the step size which would be determined more by the nature of the behaviour than the rate of attainment.

A second implication of Cunningham's theory is that when plateaux occur they result from assessed behaviours which require to be broken down into a better defined set of sub-items for the DS infant. Because DS infants require longer periods of time to consolidate such behaviours and to produce qualitatively more skilled performances, Cunningham suggests in turn that DS is associated with some impairment in the processes by which new learning is consolidated.

As stated above, Cunningham suggests that these emerging difficulties in visuo-motor coordination and sustained attention, predict the later emergence of plateaux. At the stage at which plateaux were most prevalent (i.e. around raw scores 80 - 105), Cunningham reported highest failure rates for items which demand high levels of fine motor coordination (such as placing pegs and building towers with cubes). He similarly found items requiring sustained attention or an increase in attention span (such as placing 3 cubes in a cup (item 100) or attempting to attain a third cube (item 82)) to be among those most frequently failed at this level.

As will be seen in the following chapter several of the items for which Cunningham reported high percentage failure rates among 'plateauing' subjects were found in the present study to be among those on which performance was most frequently seen to change over closely-spaced testing sessions. It will be recalled that low item-item agreement figures between test and retest protocols obtained at the same age level were to a large degree attributable to motivational factors. Often children demonstrated themselves to be capable of passing specific items on one occasion, yet withheld production of that same behaviour when it was tested on another occasion. Such failures by default were equally likely to occur in either of the two sessions. This may suggest that many of the failures which Cunningham observed in single testing sessions may not have been genuine failures at all but were the result of Ss refusing to perform to full competence.

If this explanation is accepted, it has two implications. The first relates to the consolidation of these behaviours. Cunningham suggests that the processes by which they are consolidated are impaired in DS. To support this he reports, firstly that certain behaviours take longer in DS to achieve the quality normally associated with the acquisition of these behaviours; and secondly, that these items are frequently failed. By comparison, results from the present study reveal that many of these items are not simply being frequently failed due to a genuine inability to produce the required behaviours, but that these behaviours are being withheld. Failures to produce behaviours even after they have been acquired would also suggest difficulties with consolidation, but in this explanation a motivational component is incorporated. Rather than simply taking longer to consolidate behaviours due to some impairment in the physiological processes associated with consolidation, it would appear that DS children are providing themselves with further obstacles. If skills are being unreliably produced it would not seen unreasonable to suggest that, as a consequence, they might take longer to consolidate. It may, however, be the case that a high level of failure experienced at the earlier stage in development, during the acquisition of similar but lower level skills, could explain the avoidance seen at later age levels.

A second implication relates to the overall decline in rate of intelligence. Where failures by default are recorded as genuine failures, test scores necessarily reflect a below-optimal estimate of a child's true level of cognitive functioning. Exaggeratedly low scores, particularly when they are recorded at periods when development is slowing, would imply a much sharper decline than may actually be occurring. In this study where individual Ss' optimal scores were compared with scores attained from individual sessions, this would appear to be the case: the magnitude of MDI decline was substantially reduced when 'fuller' performances, measured over two sessions, were considered. This held true even in cases where, despite the use of optimal scores, Ss did appear to have reached genuine plateaux in development.

An even more complex picture, however, emerged when these optimal score plateaux were investigated. It was revealed that the plateauing effect was in part attributable to the loss of items from performance repertoires. An average of 5 raw score points was being dropped between ages where, on the basis of optimal scores, plateaux were recorded. To maintain scores at equivalent levels between age intervals therefore, 5 new items must have been attained over the same period. At most MA levels, a gain of 5 raw score points corresponds to a 1.5 month increase in mental age level. If, as has previously been claimed (Berry et al 1984; Rauh et al in press), development in DS proceeds at approximately half the normal rate, in cases where tests were administered at 3 monthly intervals, MA would be expected to increase by 1.5 months between these intervals. In other words, if these items were not being dropped, rather than appearing to plateau, development would maintain a constant upward rate.

When group curves were amended to include *all* items previously passed, however, there was little evidence to suggest that plateaux and the frequency with which items were lost from performance repertoires could explain the overall decline in developmental rate. This is perhaps not surprising for the following 3 reasons. Firstly, plateaux were not in evidence in all cases. Secondly, as has already been demonstrated, development in DS cannot be described as progressing at a constant rate; even when children were making progress, the extent of this progress varied between each age interval. Thirdly, although where plateaux did occur, an average of 5 raw score points were being lost over age intervals, this figure obscures a considerable degree of variation among individual cases. There were cases in which, even when score losses were accounted for, developmental rate still appeared to have declined. Just as it appears fruitless to attempt to describe the course of development in DS in terms of a single mathematical function, it would appear to be equally fruitless to focus on a single explanation for the decline in developmental rate - which may itself be less characteristic of development in DS than has previously been believed.

It could moreover be suggested that even in these cases where developmental rate still appeared to have declined even this decline may not in fact have been genuine. There is little reason to suppose that these apparent losses of skills are any less likely to be examples of childrens' refusals to perform to full competence than were unreliable performances observed over more closely-spaced intervals. Indeed the same could be said of optimal scores - why should children withhold demonstration of certain behaviours on one occasion only? It often in fact appeared to be the case that children were refusing to engage in the same task in both sessions.

If the above points are accepted, perhaps any attempt to assess the cognitive functioning of children with DS will always be fraught with difficulties relating to unreliable performance. Whether attributable to motivational problems, to an impairment in the processes of consolidation or to a complex interaction between the two, clearly this lack of reliability invalidates any attempt to devise any model of development in DS on the basis of performance on psychometric tests. Poor reliability may, however, provide some insight into the developmental process in this particular population. If, as it would seem, skills are lost from DS children's repertoires, it would seem prudent firstly, to investigate the cause of their loss, and secondly, to explore any possible effects this may have on the process of development itself. The next chapter re-examines the data from the longitudinal study just described with these two aims.

CHAPTER 6:

LONGITUDINAL PERFORMANCE ON THE BSID: THE RELATIONSHIP BETWEEN UNRELIABILITY AND INSTABILITY

Findings presented in study 4 revealed a number of shortcomings inherent in the practice of devising models of development in DS children on the basis of scores from psychometric tests. In addition to problems relating to validity and reliability which were discussed in study 3, it emerged that a further issue - developmental instability - had also been implicitly overlooked by researchers relying on this method of assessing cognitive functioning in this population of children. Item-item comparisons of BSID protocols obtained from the longitudinal study revealed that DS children frequently failed to reproduce passes attained at earlier age levels. In addition the high degree of discrepancy between pass and fail patterns produced in pairs of closely-spaced testing sessions lent further support to the findings presented in study 3, with test-retest reliability again proving to be far lower than would be expected on the basis of normative studies with non-handicapped children. It was concluded that unless these issues of reliability and stability are fully considered, any theory of development in DS constructed solely from psychometric data is likely to be be inaccurate.

Although these findings have very negative implications in relation to the use of psychometric methods of assessment with DS children, at the same time they may provide useful insights into the development of cognition in this group of children. DS children appeared to be losing skills over adjacent age levels suggesting that their development does not follow the steady, incremental pattern seen in normally developing children. Given that unreliability was also found to be characteristic of DS test performance this would seem an obvious starting point for investigating the possible cause(s) of these apparent losses of skills. This chapter aims therefore to explore any relationship that may exist between unreliable performance on certain tasks at given ages and the apparent loss of ability to succeed on these same tasks at later stages in development.

Loss of skills: implications for the developmental process as a whole

It would appear that failure to perform to full competence is not simply characteristic of the behaviour of DS children in testing situations. According to the parents of many of the children who took part in the studies reported thus far, refusal to perform on demand is equally prevalent in other situations. Often when asked about their children's usual responses to being asked to demonstrate specific skills parents would give replies such as "sometimes - when it suits him" or "only when she feels like it".

As already stated, by nature of their handicapping condition, DS children are already disadvantaged in relation to learning; failure to reliably produce those skills that they do acquire can only add to this disadvantage. Although it has already been argued that development in DS cannot accurately be described simply as being slow in relation to normal development it cannot nevertheless be denied that DS children do take longer than their non-handicapped peers to reach each individual developmental milestone. This process can only be further prolonged if new acquisitions are not regularly rehearsed and consolidated, allowing the child to build on these consolidated skills and to advance to more complex activities.

Relatedly and perhaps more importantly, if the outcome of unreliable production of skills is the eventual loss of these skills, this could imply that acquisitions made subsequent to ones which have disappeared would not be fully understood. Since the work of Piaget it has become widely accepted that development proceeds as a series of stages; a full grasp of the skills and concepts relevant to each stage being essential before the next can be attained. Morss (1980) has reported that it often appears that DS learning is 'incomplete'. Where successes do occur these seem to be more of the nature of one-off achievements rather than reflecting the acquisition of any general understanding. It may well be that this lack of understanding is the cumulative result of DS childrens' failures to cement the building blocks which they have used to advance onto higher-level skills. If this is the case, clearly a predominant focus of early intervention should be to ensure that lower level behaviours are fully consolidated before children are encouraged to advance onto more complex activities.

Loss of skills: are losses genuine or do they reflect the inadequacy of the BSID?

The aim of this chapter is to determine whether any relationship can be found between unreliable performance and longitudinal instability in BSID profiles. Before attempting to speculate on the nature of any such relationship it seems necessary to examine the possibility that unstable performance - the apparent 'loss' of skills at later ages - may simply be characteristic of DS behaviour in testing situations. Parents' reports would seem to imply that unreliability is as prevalent outside the testing room as it is during assessments. As mentioned in the previous chapter, however, it cannot be assumed that when a child fails an item passed at an earlier age he/she has necessarily lost the requisite skills for passing that item. Over closely-spaced intervals where performance on certain items was seen to change this change often appeared to result from the child's failure to engage in the task on one of the two occasions. Many failures could not therefore be confidently attributed to any straightforward absence of the required levels of cognitive ability. The apparent disappearance of skills over 3 and 6 monthly intervals may simply be explained in the same way: rather than indicating difficulties with consolidation, consistent failures at later ages may have simply resulted from refusals to perform to full competence on certain tasks on both testing occasions. Rather than having any wider implications in relation to the learning process it may be that an explanation for the frequency with which children with DS were found to drop test items may be provided in the nature of the tests themselves. As was discussed in study 3, for example, the items included in the BSID may simply be inappropriate for use with this - or indeed any - group of mentally handicapped children. A DS child of 5 years of age may be functioning at a level similar to that of a non-handicapped child of half that age but this does not mean that he/she has not outgrown the sorts of tasks and objects which are stimulating for the younger child.

It was also mentioned in study 3 that, for similar reasons, a child may fail to produce the behaviour required for passing a given item because he/she has gone onto a more constructive or complex activity with the objects presented. In the BSID manual, examiners are advised to consider such failures when setting basal levels. The example which is given to illustrate this point involves the 1 inch cubes. It is stated that a child may not demonstrate him/herself capable of, for example, retaining 2 cubes because to that child the cubes may provide an opportunity to demonstrate a more advanced skill (e.g. building a tower with the cubes). In this case it is stated that the child 'surpasses' the item rather than passes or fails it and should receive a credit.

This study will re-examine the longitudinal BSID protocols in an attempt to determine why it is that items passed at earlier ages are not reproduced after 3 and 6 monthly intervals. In particular this investigation will focus on evidence of unreliability at earlier developmental stages as a possible cause of this apparent loss of skills at later ages. Other, more 'practical' reasons such as lack of suitability of test items and the possibility that children have surpassed certain items will also be considered, however.

STUDY 5

METHOD

Data from the longitudinal BSID protocols were re-examined in two separate analyses; an analysis of short-term test-retest reliability and an analysis of the longer term stability of performance profiles.

A/ Test-retest reliability

As in the study reported in Chapter 4, protocols were investigated for any variation in pass/fail patterns produced in the two closely-spaced sessions administered at each age level. A child's performance on any item was considered *unreliable* if s/he had passed that item in one of the two closely-spaced sessions but had failed it in the other.

B/ Long-term stability

Longitudinal performance was investigated for evidence of item loss. An item was considered 'lost' if it had been passed in at least one of the two closely-spaced sessions at any age level but was failed in both closely-spaced sessions at a later age level. In cases where items were lost from protocols, performance on those items was considered to be *unstable*.

As already mentioned, examiners using the BSID are alerted to the possibility that children may surpass certain low-level test items. It is stated in the test manual that rather than indicating the absence of the skills required to pass such items, failures may indicate that the child has advanced onto higher-level activities with the test materials. In the present study, given the frequency with which performance was seen to change both in the long-term and in the short-term and on both high- and low-level items, it was impossible to determine with any confidence whether each individual case of item loss could be attributed to the subject having surpassed that particular item. In scoring test performance it was therefore decided only to credit items in each session if the requisite skill to pass had actually been demonstrated during that session.

RESULTS

A/ Test-retest reliability

Table 6:1 contrasts the mean percentage agreement figures for individual BSID items with, where possible, the reliability data for these same items from the Werner and Bayley (1966) and the Horner (1980) BSID reliability studies (see also p. 85-86).

Table 6:1 Test-Retest percentage agreement figures for each item

				% test-retest agreement			
	Item	NT-R	NU	DS S's	H S's 9mths	H S's 15mths	W & B S's
10:	eyes follow moving person	2	-	100	nr	nr	nr
11:	responds to voice	1	-	100	nr	nr	nr
12:	vertical eye coordination:light	1	-	100	nr	nr	nr
13:	vocalises once or twice	1	-	100	nr	nr	nr
14:	vertical eye coordination: red ring	1	-	100	nr	nr	nr
15:	circular eye coordination: light	1	-	100	nr	nr	nr
16:	circular eye coordination: red ring	3	-	100	nr	nr	nr
17:	free inspection of surroundings	5	2	60	nr	nr	nr
18:	social smile: E talks and smiles	5	3	40	nr	nr	nr
19:	turns eyes to red ring	5	-	100	nr	nr	nr
20:	turns eyes to light	5	-	100	nr	nr	nr
21:	vocalises at least 4 times	5	-	100	nr	nr	nr
22:	anticipatory excitement	5	-	100	nr	nr	nr
23:	reacts to paper on face	5	-	100	nr	nr	nr
24:	blinks at shadow of hand	5	-	100	nr	nr	nr
25:	visually recognises mother	5	-	100	nr	nr	nr
26:	social smile: E smiles, quiet	5	-	100	nr	nr	nr
27:	vocalises to E's social smile and talk	6	-	100	nr	nr	nr
28:	searches with eyes for sound	6	3	50	nr	nr	nr
29:	eyes follow pencil	6	-	100	nr	nr	nr
30:	vocalises 2 different sounds	6	-	100	nr	nr	nr
31:	reacts to disappearance of face	7	-	100	nr	nr	nr
32:	regards cube	8	-	100	nr	nr	nr
33:	manipulates red ring	8	4	50	nr	nr	nr

NT-R => Number of Test-Retest presentations; NU=> Number of times unreliable;
DS S's => DS Subjects; H S's Horner Subjects; W & B S's => Werner & Bayley Subjects;
nr => figures which were nor reported in the Horner and the Werner & Bayley studies.

	Item	NT-R	NU	DS S's	H S's 9mths	H S's 15mths	W & B S's
					% test-retest agreement		
34:	glances from one object to another	8	2	75	nr	nr	nr
35:	anticipatory adjustment to lifting	8	-	100	nr	nr	nr
36:	simple play with rattle	8	1	87.5	nr	nr	nr
37:	reaches for dangling ring	7	1	85.8	nr	nr	nr
38:	follows ball visually across table	10	-	100	nr	nr	nr
39:	fingers hand in play	10	5	50	nr	nr	nr
40:	head follows dangling ring	11	-	100	nr	nr	nr
41:	head follows vanishing spoon	10	1	90	nr	nr	nr
42:	aware of strange situation	10	1	90	nr	nr	nr
43:	manipulates table edge slightly	11	1	80.9	nr	nr	nr
44:	carries ring to mouth	13	4	69.3	nr	nr	nr
45:	inspects own hands	13	4	69.3	nr	nr	nr
46:	closes on dangling ring	11	2	81.82	nr	nr	nr
47:	turns head to sound of bell	12	1	91.7	nr	nr	nr
48:	turns head to sound of rattle	11	2	81.82	nr	nr	nr
49:	reaches for cube	8	1	87.5	nr	nr	nr
50:	manipulates table edge actively	8	1	87.5	nr	nr	nr
51:	eye-hand coordination in reaching	9	1	88.9	nr	nr	nr
52.	regards pellet	9	3	66.6	nr	nr	nr
53:	mirror image approach	9	4	55.6		nr	95
54:	picks up cube	11	3	72.8	nr	nr	nr
55:	vocalises attitudes	17	2	88.3		nr	73
56:	retains 2 cubes	16	7	56.25	nr	nr	nr
57:	exploitive paper play	15	1	93.3	nr	nr	nr
58:	discriminates strangers	16	-	100		nr	59
59:	recovers rattle in crib	19	1	94.8	nr	nr	nr
60:	reaches persistently	19	4	79	nr	nr	95
61:	likes frolic play	21	-	100	nr	nr	58
62:	turns head after fallen spoon	21	8	62	nr	nr	95
63:	lifts inverted cup	23	5	78.3	nr	nr	86
64:	reaches for 2nd cube	24	10	58.4	nr	nr	95
65:	smiles at mirror image	25	5	80	nr	nr	nr

	Item	NT-R	NU	% test-retest agreement DS S's	H S's 9mths	H S's 15mths	W & B S's
66:	bangs in play	24	3	87.5	nr	nr	95
67:	sustained inspection of red ring	24	8	66.6	nr	nr	86
69:	transfers object hand to hand	25	2	92	nr	nr	95
70:	picks up cube deftly and directly	26	5	80.8	100	-	91
71:	pulls string: secures ring	25	3	88	100	-	nr
72:	interest in sound production	26	3	88.5	100	-	nr
73:	lifts cup with handle	26	7	73.1	92	-	89
74:	attends to scribbling	29	4	86.3	100	-	nr
75:	looks for fallen spoon	28	11	60.8	92	-	85
76:	playful response to mirror	31	4	81.1	71	-	nr
77:	retains 2 of 3 cubes offered	39	16	59	92	-	95
78:	manipulates bell: interest in detail	38	10	73.7	92	-	nr
79:	vocalises 4 different syllables	39	3	92.4	71	-	67
80:	pulls string adaptively: secures ring	39	6	84.7	75	-	nr
81:	cooperates in games	43	8	81.4	71	-	58
82:	attempts to secure 3 cubes	45	16	64.5	71	-	nr
83:	rings bell purposively	45	9	80	75	-	nr
84	listens selectively to familiar words	42	1	97.7	83	-	41
85:	says "dada" or equivalent	44	10	79.3	54	-	51
86:	uncovers toy	44	13	70.5	92	-	85
87:	fingers holes in peg board	43	7	83.8	83	-	nr
88:	picks up cup: secures cube	44	13	70.5	92	-	nr
89:	responds to verbal request	44	5	88.7	62	-	nr
90:	puts cube in cup on command	45	9	80	70	-	70
91:	looks for contents of box	45	7	84.4	88	-	85
92:	stirs with spoon in imitation	45	6	86.7	75	-	nr
93:	looks at pictures in book	48	6	81.5	100	-	nr
94:	inhibits on command	42	4	90.5	88	-	72
95:	attempts to imitate scribble	43	7	83.8	92	-	96
96:	unwraps cube	42	8	81	83	-	85
97:	repeats performance laughed at	42	6	85.8	71	-	54

	Item	NT-R	NU	% test-retest agreement DS S's	H S's 9mths	H S's 15mths	W & B S's
98:	holds crayon adaptively	42	5	88.9	100	-	92
99:	pushes car along	42	9	78.6	96	-	nr
100:	puts 3 or more cubes in cup	42	6	85.8	88	100	nr
101:	jabbers expressively	44	14	66.6	67	88	nr
102:	uncovers blue box	45	10	77.8	67	96	nr
103:	turns pages of book	46	5	89.2	75	100	nr
104:	pats whistle doll, in imitation	46	11	76.1	62	79	91
105:	dangles ring by string	42	10	76.2	71	96	nr
106:	imitates words	29	7	75.9	88	46	88
107:	puts beads in box	28	14	50	96	100	nr
108	places 1 peg prepeatedly	29	4	86.2	92	96	nr
109:	removes pellet from bottle	30	11	63.4	100	100	nr
110:	blue board: places 1 round block	32	9	71.9	-	100	nr
111:	builds tower of 2 cubes	38	18	52.7	-	96	nr
112:	spontaneous scribble	37	11	70.3	-	88	86
113:	says 2 words	33	2	94.4	-	54	91
114:	puts 9 cubes in cup	35	9	74.3	-	92	nr
115:	closes round box	35	12	66	-	100	nr
116:	uses gestures to make wants known	38	5	87	-	100	nr
117:	shows shoes	38	16	58	-	100	nr
118:	pegs placed in 70 secs	34	6	82.4	-	71	nr
119:	builds tower of 3 cubes	39	7	82.1	-	54	nr
120:	pink board: places round block	35	11	68.6	-	92	nr
121:	blue board: places 2 round blocks	34	8	72.5	-	75	nr
122:	attains toy with stick	35	15	58	-	83	nr
123:	pegs placed in 42 secs	38	4	89.5	-	42	nr
124:	names 1 object	33	5	84.8	-	33	nr
125:	imitates crayon stroke	33	11	67	-	75	nr
126:	follows directions, doll	34	5	85.3	-	79	nr
127:	uses words to make wants known	34	3	91.2	-	92	nr
128:	points to parts of doll	32	3	90.7	-	96	nr
129:	blue board: places 2 round and 2 square blocks	35	8	77.2	-	88	nr
130	names 1 picture	39	8	79.5	-	79	nr
131:	finds 2 objects	39	12	71.4	-	67	nr

	Item	NT-R	NU	% test-retest agreement DS S's	H S's 9mths	H S's 15mths	W & B S's
132:	points to 3 pictures	36	10	69.3	-	88	nr
134:	pegs placed in 30 secs	35	6	86.2	-	75	nr
135:	differentiates scribble from stroke	35	12	82.6	-	83	nr
136:	sentence of 2 words	38	6	65.8	-	96	nr
137:	pink board: completes	35	6	84.3	-	67	nr
138:	names 2 objects	34	10	82.9	-	88	nr
139:	points to 5 pictures	34	9	70.6	-	92	nr
140:	broken doll: mends approximately	33	5	73.6	-	67	nr
141:	names 3 pictures	33	5	84.9	-	96	nr
142:	blue board: places 6 blocks	38	7	84.9	-	71	nr
143:	builds tower of 6 cubes	34	15	55.9	-	96	nr
144:	discriminates 2: cup, plate, box	32	11	65.7	-	100	nr
145:	names watch, 4th picture	33	7	78.8	-	100	nr
146:	names 3 objects	33	9	72.8	nr	-	nr
147:	imitates strokes: vertical and horizontal	33	12	63.7	nr	-	nr
148:	points to 7 pictures	30	5	83.4	nr	-	nr
149:	names 5 pictures	36	2	94.5	nr	-	nr
150:	names watch, 2nd picture	30	9	70	nr	-	nr
151:	pink board: reversed	27	5	81.5	nr	-	nr
152:	discriminates 3: cup, plate, box	27	8	70.4	nr	-	nr
153:	broken doll: mends exactly	29	6	79.4	nr	-	nr
154:	train of cubes	29	6	79.4	nr	-	nr
155:	blue board: completes in 150 secs	29	6	79.4	nr	-	nr
156:	pegs placed in 22 seconds	30	10	66.6	nr	-	nr
157:	folds paper	30	9	70	nr	-	nr
158:	understands 2 prepositions	28	7	75	nr	-	nr
159:	blue board: completes in 90 secs	29	7	75.8	nr	-	nr
160:	blue board: completes in 60 secs	27	7	74.1	nr	-	nr
161:	builds tower of 8 cubes	18	9	50	nr	-	nr
162:	concept of one	11	6	45.5	nr	-	nr
163:	understands 3 prepositions	4	1	75	nr	-	nr

NT-R => Number of Test-Retest presentations; NU=> Number of times unreliable;
DS S's => DS Subjects; H S's Horner Subjects; W & B S's => Werner & Bayley Subjects;
nr => figures which were nor reported in the Horner and the Werner & Bayley studies.

For the total range of 154 items presented to Ss in this study (items 10 - 163), test-retest agreement averages 79.6%. This overall reliability figure, higher than that found in the smaller test-retest study presented in Chapter 4, is also higher than that found for the small subsection of items tested for reliability in the Werner and Bayley study (1966). As has already been pointed out, however, (p.91) the low level of reliability found in that study - 76.4% - is likely to have been influenced by the particular items tested (many of which were subsequently dropped). When compared to the reliability demonstrated by Horner's more recent sample of 9 and 15 month old non-handicapped children, DS performance proved to be considerably less reliable. Mean percentage agreement for the set of items (70 - 109) presented to the 9 month olds in the Horner study was 84.8% For the same set of items the mean figure found in this study was 79.33% Similarly, the percentage agreements on items presented to Horner's 15 month sample (items 100-145) was 76.06% in this study in comparison to Horner's 84.5%. T-test comparison of item-item agreement figures for the entire set of items used in the Horner study (items 70 -145) with DS agreement figures obtained for the same items revealed a significant difference ($t = 3.626$, df 75, $p < 0.0005$).

Table 6:2 presents the group of items on which performance was found to be most unreliable. As was found in the study presented in Chapter 4 DS performance proved to be particularly variable on items testing cube behaviours and those involving the crayon and paper.

Table 6:2 Items showing poor reliability over closely spaced testing sessions.

% Agreement less than 60			% Agreement less than 70		
53:	mirror image approach	55.6	44:	carries ring up to mouth	69.3
56:	retains 2 cubes	56.25	45:	inspects hands	69.3
64:	reaches for 2nds cube	58.4	52:	regards pellet	66.6
77:	retains 2/3 cubes offered	59	62:	turns head after fallen spoon	62
107:	puts beads in box	50	67:	sustained inspection of ring	66.6
111:	builds tower of 2 cubes	52.7	75:	lifts cup with handle	60.8
117:	shows shoes	58	82:	attempts to secure 3 cubes	64.5
122:	attains toy with stick	58	101:	jabbers expressively	66.6
143:	builds tower of 6 cubes	55.9	109:	removes pellet from bottle	63.4
161:	builds tower of 8 cubes	50	115:	closes round box	66
162:	concept of 1	45.5	120:	pink board: places round block	68.6
			125:	imitates crayon stroke	67
			131:	finds 2 objects	69.3
			135:	differentiates scribble from stroke	65.8
			144:	discriminates 2: cup, plate, bix	65.7
			147:	imitates strokes: vertical & horizontal	63.7
			156:	peg placed in 22 secs.	66.6

Failures by default

As in the study reported in Chapter 4, test-retest protocols were examined to determine the extent to which unreliable performances may have been attributable to failures by default. Again it was found that a substantial proportion of cases of item variability were attributable to children's failures to engage in the item in one of the two sessions. A total of 853 occurrences of performance change were recorded. Ninety-two such cases were omitted from the analysis below on the grounds that it was not possible to distinguish between criterion fails and fails by default (see p). In the 761 remaining cases of variability 487 - 64% - were clearly due to Ss' refusal to engage in the tasks.

Test-retest reliability measures can only identify items on which children underperformed on one of two occasions. In many cases, however, children were seen to avoid (and consequently fail) the same item in both sessions. Because of such cases in which performance on a given item does not change, the resulting high test-retest 'reliability' figure for a particular task may obscure the fact that performance may not necessarily be reliably reflecting children's true competence.

Table 6:3 shows the items which were most often failed by default in either or in both sessions. It can be seen that although the majority of these items obtained low test-retest reliability figures, many were also frequently avoided on both testing sessions thereby artefactually inflating the reliability figures obtained. For example, only 3 of the 32 Ss who were presented with item 128 (points to parts of doll) changed their performance on this item giving it a high percentage agreement figure of 90.7. All 3 failed to engage in this item on one of the two occasions. A further 12 Ss however failed this item by default in both sessions. Item 130 (names 1 picture) also already showing a comparatively high reliability figure - 79.5 - was failed twice by an additional 8 Ss who refused to engage in this task when it was presented to them, suggesting an even lower reliability figure for this item.

Other, less extreme examples can also be seen in Table 6:3. Item 107 (puts beads in box), for example, was avoided and consequently failed by 57.14% of Ss presented with this task. Other items frequently failed by default included items 111 (builds tower of 2 cubes), item 131 (finds two objects), and item 144 (discriminates 2: cup, plate, box). In addition a large proportion of items avoided in common by Ss were those testing cube behaviours and discrimination skills.

Table 6:3 Items most frequently failed by default

Item		In one session	In both sessions	% Test-Retest Agreement
64:	reaches for 2nd cube	6	2	58.4
77:	retains 2/3 cubes offered	9	3	59
82:	attempts to secure 2nd cube	10	6	64.5
107:	puts beads in box	13	3	50
110:	blue board: 1 round block	4	3	71.9
111:	builds tower of 2 cubes	14	4	52.7
117:	shows shoes	7	4	58
120:	pink board: places round block	7	2	68.6
128:	points to parts of doll	3	12	90.7
130:	names 1 picture	3	8	79.5
131:	finds 2 objects	12	8	71.4
132:	points to 3 pictures	8	3	69.3
133:	broken doll: mends marginally	5	10	72.3
135:	differentiates scribble from stroke	7	2	82.6
143:	builds tower of 6 cubes	7	4	55.9
144:	discriminates 2; cup, plate, box	8	6	65.7
152:	discriminates 3: cup, plate, box	2	4	70.4
154:	train of cubes	3	4	79.4
161:	tower of cubes	5	1	50

B/ Long-term stability of performance profiles

Items lost over adjacent pairs of ages

By contrast, failure to engage was found to account for only 16% of instances of item loss. Of the total of 268 items which were dropped from Ss' protocols over age levels, only 44 were failed by default in either or both testing sessions administered at the age level at which they were found to have disappeared.

Table 6:4 shows the items which most frequently disappeared from childrens' protocols over adjacent ages. It can be seen that few of these items also attained low reliability figures. This does not necessarily suggest, however, that there is little association between items which were both produced unreliably and those which were dropped over ages - in fact it may suggest the opposite. For this study reliability figures were calculated for each item from the total number of cases in which that item was presented to any child on any occasion. As most items were presented to the same children at more than one age level this means that if a child's performance was unreliable on a specific item at one age level but was reliably failed at the next, the item-item disagreement found at the earlier age level would be cancelled out by the agreement at the next age level. As will be seen below, however, there were in reality few items on which Ss did in fact produce both unreliable and unstable performances in common. Overall reliability figures for individual items were not therefore affected to any great extent by this effect. The fact that there was a high degree of agreement between reliability figures obtained in this study and those reported in the cross-sectional reliability study presented in Chapter 4 suggests that this effect, even if present, would be weak.

Although in the minority, there were several items on which performance was particularly variable in closely-spaced sessions and which were also frequently dropped from protocols over adjacent age levels. This correspondence was strongest on items testing cube behaviours (item 77: retains 2/3 cubes; item 82 :attempts to secure 3 cubes: item 161: builds tower of 8 cubes) and on item 117 (shows shoes).This suggests that these particular items may have been especially problematic for this group of children.

Table 6:4 Items most frequently dropped over age levels

Item		No. cases dropped over at least 1 age level	No. cases dropped over more than 1 age level	% Test-Retest agreement
73:	lifts cup with handle	4	2	73.1
77:	retains 2/3 cubes offered	4	6	59
78:	manipulates bell - interest in detail	8	5	73.7
81:	cooperates in games	5	2	81.4
82:	attempts to secure 3 cubes	6	2	64.5
85:	say "dada" or equivalent	6	-	77.3
86:	uncovers toy	5	5	70.5
88:	picks up cup: secures cube	4	1	70.5
92:	stirs with spoon in imitation	4	1	86.7
94:	inhibits on command	6	5	90.5
96:	unwraps cube	4	-	81
105:	dangles ring	5	-	76.2
117:	shows shoes	5	2	58
130:	names 1 picture	5	-	79.5
133:	broken doll: mends marginally	4	2	86.2
161:	tower of 8 cubes	4	1	50

Group data: items showing low reliability/high instability

Item 82 was also frequently dropped from children's performance repertoires in the Cunningham (1979) study. Bayley, however, specifically identified low level cube behaviour items of this sort to be ones which children may fail because they have surpassed them i.e have advanced onto more complex activity. It was therefore decided to investigate the 4 cases in which item 77 was dropped from Ss' protocols and the 6 cases in which item 82 was dropped for evidence that children had in fact developed more advanced cube skills.

Item 77: retains 2/3 cubes

This item was dropped by S4 at 9 months, S12 at 12 months and Ss 1 and 8 at 15 months. The performance of Ss 1,4 and 8 on this item had varied over closely-spaced sessions administered at the preceding age level. Both Ss 1 and 4 had passed it in only one of the two sessions; Subject 8, although gaining a credit for this item in both sessions produced a more sophisticated response in session 1 (see case study 1 p). Subject 12 was only tested once at 9 months. It is not therefore possible to identify any link between unreliable and unstable performance in this particular case. Nevertheless, over the entire subsequent 15 months during which he participated in this study Subject 12 did not once reproduce this item. It was regained in subsequent sessions by Subjects 1 and 4 but was not reliably produced over closely-spaced tests in either case. In fact it again disappeared and reappeared from Subject 4's protocol between 18 and 24 months. Only Subject 8 reliably produced this item on 2 occasions 3 months after it had first dropped out.

This item remained absent from the protocols of Ss 4 and 12 for 6 and 9 months respectively with no evidence that any higher level cube items had been attained during these periods. In fact, in both cases, even lower level items in this series were also either dropped altogether or were failed in 1 of 2 sessions presented at the same age level (item 49: reaches for cube; item 54: picks up cube; item 56: retains 2 cubes; item 64: reaches for 2nd cube). Both Ss did eventually gain higher level items at much later ages (Subject 4: 15 months - item 90: puts cube in cup; item 82: attempts to secure 3 cubes; 21 months - item 100: puts 3 or more cubes in cup; Subject 12: 21 months - item 90: 24 months - item 100) but in every case these items were produced unreliably at these age levels.

Item 82: attempts to secure 3 cubes

Three Ss who had dropped item 77 also dropped this item (Subject 8 at 15 months; Subject 1 at 15 months; Subject 4 at 24 months) Item 82 also disappeared from the protocols of Ss 5 (15 months), 9 (18 months) and 10 (15 months). Ss 1,4 and 5 had passed this item in only one of the sessions administered 3 months prior to its disappearance from their protocols.

In only one case (Subject 5) was there no evidence that any higher level cube item had been attained. Subjects 8 and 10 both passed item 100 (puts 3 or more cubes in cup) for the first time at 15 months, although that particular item again became unreliable for Subject 10 at 18 months when item 82 actually reappeared - but was unreliably produced. It was dropped once more from this S's protocol at 21 months although items 111 (builds tower of 2 cubes) and 114 (puts 9 cubes in cup) were attained - albeit unreliably - at this age level. Item 111 remained unreliable until 27 months for this subject. Subject 8 also gained item 111 at 18 months although it was not produced reliably either at this age or at 21 months. Similarly, every higher level cube item gained by Ss 1,4 and 9 was unreliably produced at the age level at which item 82 was lost.

More qualitative analyses of individual S's performance on these cube items will be presented in the case studies below.

Item 117: shows shoes

This item was also found to have a low test-retest percentage agreement figure and to be among the group of items which most frequently disappeared from childrens' protocols between age levels. This item is the earliest discrimination item to appear in the BSID.

Five children dropped item 117 after having passed it at an earlier age level. Subject 10 failed it twice at both 24 and 27 months, having passed it in one of the two sessions presented at 21 months. This item was also passed (once) by Subject 14 at 21 months and in both sessions at 27 months but she failed twice at 24, 33 and 36 months. Subject 18 passed it in one of each of the two sessions at 36 and at 48 months, failing it twice at 42 months and again in both sessions at 54 months. Ss 21 and 22 lost this item only once at 42 and 54 months respectively; in both cases it was regained at the next age level although Subject 21 failed it in the second session.

Subjects 10 and 14 had attained this item for the first time at the age level preceding that in which it had first dropped out. In both cases it was only passed in one of the two closely-spaced testing sessions but in neither case could failure clearly be attributed to failure to engage. In fact both subjects seemed genuinely unable to pass this item on each of these occasions. Failure to engage occurred at later ages: at 27 months for Subject 10 and at 33 and 36 months for Subject 14. Subject 18 failed to engage in this task in the second session at 36 months (the age level at which she entered this study), but passed it in session 1 at this age. She did not appear to fail this item by default in any subsequent session, however, and genuinely seemed to have lost the ability to pass this task. Subject 21 refused to engage in item 117 in both sessions at 21 months but this did not seem to be the case for Subject 22 who, at 54 months, seemed genuinely unable to provide the correct response to E's request to point to his shoes.

It is unlikely that failures to engage on this item could be attributable to any subject having surpassed it - this is not the sort of behaviour that can really be surpassed. The protocols of these 5 children were therefore examined in an attempt to provide an explanation as to why this item was dropped in all 5 cases.

This investigation revealed that when item 117 was dropped this generally occurred when Ss had attained higher level discrimination items for the first time, or were avoiding these more difficult items. Although there was no evidence that either Subjects 10 or 14 had attained any higher level discrimination skill when item 117 was first dropped, both Ss were seen to avoid other discrimination items when this item disappeared for the second time. When presented with the picture cards in the first of two sessions administered at 27 months, Subject 10 refused to engage in this task. She did however cooperate in the second session and was credited with items 132 and 139 (points to 3/5 pictures). In contrast, Subject 14 refused to engage in every discrimination item she was presented with at 33 and 36 months (item 128: points to parts of doll; item 130: names 1 picture; item 132: points to 3 pictures). The same items were also avoided in both closely-spaced sessions by Subject 18 when she first dropped item 117 (42 months) which then reappeared at 48 months when item 132 was passed (once) for the first time. The picture items 130 and 132 were unreliably passed by this S at 54 months when item 117 again dropped out of her protocol. Items 128 (points to parts of doll) and 144 (discriminates 2: cup, plate, box) were also avoided in both sessions at this age level.

When item 117 was first dropped by Subject 21 at 42 months, he also refused to engage in every discrimination item with which he was presented. At 48 months he again refused some of the same items - including item 117 - although items 144 (discriminates 2: cup, plate, box) and 146 (names 3 objects) were produced in the second of the two sessions administered at this age level. Neither item was passed in the first session, however, whereas he did receive credit in this session for item 117.

At 48 months Subject 22 passed item 117 twice but refused to engage in item 128 (points to parts of doll) on either occasion. Item 144 (discriminates 2: cup, plate, box) was passed for the first time at this age level although it was avoided in the second session. It was then avoided twice at 54 months, together with items 117 and 128 (points to parts of doll) and items 145/150 (names watch 2nd/4th picture). He did however pass picture items 130 and 132 on both occasions at this age level.

Item 161; builds tower of 8 cubes

This item disappeared from the protocols of 4 subjects: Subject 17 (24 months), Subject 19 (48 months), Subject 20 (54 months) and Subject 24 (60 months). In every case it had been unreliably passed at the preceding age level.

Data from individual subject's protocols: items which were both unreliable and unstable

Although there were few *items* on which Ss were found *in common* to produce both unreliable and unstable performances there were many instances in which this link between unreliable and unstable production of specific items was in evidence in the protocols of individual Ss.

It will be recalled that, when this study commenced, several Ss were only tested on one occasion at each age level. Although several of these Ss were found to drop items over adjacent ages it was not possible to include these single session protocols in this particular analysis, the purpose of which was to explore the possible link between short-term unreliability and long-term instability. Of the total of 217 cases in which an item had dropped out over at least one age level therefore, 34 had to be omitted. (It is perhaps worth noting that given the extent to which protocols were found to differ over closely-spaced intervals, it seems highly probable that prior performance in several of these 34 cases of item loss might well have varied had Ss been tested twice at the earlier age level). This left 181 cases of item loss to be investigated for evidence of previous unreliable production. Performance at the preceding age level was found to have been unreliable in 105/181 of these cases - 58% This suggests that there may be a causal relationship between unreliable and unstable performance.

Table 6:5 presents every item for which a correspondence between unreliable and unstable performance was found. It can be seen that these items cover a wide range of different skills and that for several items the link between performance change and later disappearance was found in single cases only. To investigate this relationship between unreliability and instability in more detail the longitudinal profiles of 3 children were examined for evidence which might explain the disappearance of these skills between successive age levels.

Table 6:5 Items showing both unreliability and instability

Cube Behaviours		**No. Cases**
56:	retains cubes	2
64:	reaches for 2nd cube	2
70:	picks up cube deftly & directly	1
77:	retains 2 of 3 cubes offered	2
82:	attempts to secure 3 cubes	3
143:	builds tower of 6 cubes	2
161:	builds tower of 8 cubes	4
		16

Object Concept		No Cases
62:	turns head after fallen spoon	1
75:	looks for fallen spoon	1
86:	uncovers toy	2
88:	picks up cup: secures cube	2
91:	looks for contents of box	2
96:	unwraps toy	2
131:	finds 2 objects	1
		11

Crayon & Paper		No. Cases
74:	attends to scribbling	1
95:	attempts to imitate scribble	3
112:	spontaneous scribble	1
125:	imittes crayon stroke	3
135:	differentiates scribble from stroke	2
157:	folds paper	1
		11

Discrimination		No Cases
117:	shows shoes	3
124:	names 1 object	1
130:	names 1 picture	2
144:	discriminates 2: cup, plate, box	3
148:	points to 7 pictures	1
149:	names 5 pictures	2
150:	names watch: 2nds picture	1
152:	discriminates 3: cup, plate, box	1
		14

Imitation		No. Cases
83:	rings bell	1
87:	fingers holes in peg board	2
92:	stirs with spoon in imitation	3
99:	pushes car	1
105:	dangles ring	2
106:	imitates words	2
115:	closes round box	1
122:	attains toy with stick	3
		15

Pegs & Form Boards		No. Cases
110:	blue boards: places 1 round block	1
134:	pegs placed in 30 secs.	1
137:	pink board: completes	1
142:	blue board: places 6 blocks	2
156:	pegs placed in 22 secs	1
		6

Broken Doll		No. Cases
133:	mends marginally	1
140:	mends approximately	1
153:	mends exactly	3
		5

Speech & Comprehension		No. Cases
79:	vocalises 4 different syllables	1
85:	says dada	3
89:	responds to verbal request	1
94:	inhibits on command	1
136:	sentence of 2 words	2
158:	understands 2 prepositions	1
162:	concept of 1	2
163:	understands 3 prepositions	1
		12

Other		No. Cases
45:	fingers hands	2
52:	regards pellet	1
63:	lifts inverted cup	1
67:	sustained inspection of ring	2
72:	interest in sound production	1
73	lifts cup with handle	1
78:	manipulates bell: interest in detail	4
81:	cooperates in games	2
		15
	Total	105

Individual case studies

To give as broad a perspective as possible of the range of issues which emerged from analyses of the relation between unreliable and unstable test performance, 3 subjects exemplifying contrasting themes were selected for in-depth case study.

Subject 8 (H) was selected because the majority of items which were dropped from her protocol between adjacent age levels fell into 3 clear categories: items testing imitation skills, those testing the development of the object concept, and items involving the 1 inch cubes.

The profiles produced by Subject 12 (N) were selected for two reasons: firstly, because he became progressively more 'untestable' as he grew older; secondly because it appeared that this particular subject's difficulties were largely attributable to the severity of his motor deficit which seemed to be having a cumulative effect on his cognitive development.

The third subject - Subject 18 (S) was selected as a case study because of the contrast between the MA curve derived from her session 1 scores and that obtained when all items previously passed were included in her scores (see p.128). It will be recalled that when session 1 scores only were considered, this particular subject seemed to have made very little developmental progress over the 18 month period during which she participated in this study, remaining on a plateau for 12 months. In comparison, when her score profile was amended to include all items passed at earlier age levels, the resulting MA curve represented a pattern of constant developmental progress.

CASE STUDY 1

Subject 8 (H): 9 - 21 months

Personal details

Sex - F

Karyotype - standard trisomy 21

Health - very good: no specific health difficulties

Test details

Age level	9m	12m	15m	18m	21m
No. sessions	1	2	2	2	2

Unreliable items

12m

81 : cooperates in games
85 : says dada
90 : puts cube in cup
95 : attempts to imitate scribble
105 : dangles ring by string

15m

81 : cooperates in games
92 : stirs with spoon in imitation
96 : unwraps cube
97 : repeats performance laughed at
98 : holds crayon adaptively
101 : jabbers expressively

18m

98 : holds crayon adaptively
101 : jabbers expressively
106 : imitates words
111 : builds tower of 2 cubes
116 : uses gestures to make wants known
117 : shows shoes

21m

104 : pats whistle doll
111 : builds tower of 2 cubes
115 : closes round box
117 : shows shoes
121 : blue board : places 2 round blocks
129 : blue board : places 2 round & 2 square blocks
131 : finds 2 objects
132 : points to 3 pictures
137 : pink board : completes
142 : blue board : places 6 blocks

Unstable items

Items dropped at 15m	Item dropped at 18m
77 : retains 2/3 cubes offered	96 : unwraps cube +
82 : attempts to secure 3 cubes	
85 : says dada +	
86 : uncovers toy	Item dropped at 21m
88 : picks up cup: secures cube	106 : imitates words +
95 : attempts to imitate scribble +	
104 : pats whistle doll in imitation	

\+ *denotes items which were also unreliable at the preceding age level.*

Analysis of Performance Profiles

Between 9 and 12 months H made rapid progress, gaining a total of 14 items. Seven of these items involved imitation skills, 3 items tested object concept development, and 3 play/social items. No items which had been passed at 9 months were failed in either session at 12 months.

At 15 months however, although H attained a further 5 items she also dropped 7 items which she had passed at 12 months (she therefore reached a score plateau). No new imitation items were attained but she lost two which she had acquired at 12 months (item 95: attempts to imitate scribble; item 104: pats whistle doll in imitation). One new acquisition involved an object concept skill (item 96: unwraps cube), but this item was apparently acquired at the expense of 2 object concept items which had been attained at 12 months (item 86: uncovers toy; item 88: picks up cup: secures cube). Similarly, although item 102 (uncovers blue box) was attained at 18 months, item 96 (unwraps cube) which had been passed for the first time at 15 months then disappeared from her protocol. A number of changes in response to items testing cube behaviours was also present.

A more qualitative analysis of H's behaviour on presentation with the above items may provide some explanation for these patterns of gains and losses of imitation, object concept, and cube items.

Imitation items

Of the 7 new acquisitions made at 12 months which involved imitation-type skills, only 2 were reliably passed at this age level and at all subsequent ages (item 83: rings bell; item 87: fingers holes in peg board). At 12 months H's performance changed in a pass-to-fail direction on item 105 (dangles ring by string) and in the opposite direction on items 90 (puts cube in cup) and 95 (attempts to imitate scribble). The crayon and paper item was lost at 15 months as was item 104 (pats whistle doll in imitation) which had been reliably passed at the preceding age level. Also reliably passed at 12 months was item 92 (stirs with spoon in imitation) although this particular item was only passed in the first of the two sessions administered at 15 months.

Items 106 (imitates words) and 111 (builds tower of 2) were attained but unreliably produced at 18 months. Although item 111 was again produced in one session at 21 months, item 106 was dropped. In addition item 115 (closes round box) was passed for the first time at this age level but only in the second of the 2 sessions administered.

The next section will provide detailed descriptions of H's behaviour when presented with these imitation items on which her performance was found to change over closely-spaced sessions.

Items 90: puts cube in cup; 95: attempts to imitate scribble; 105: dangles ring by string

As stated above these 3 items were attained at 12 months but were not reliably passed at this age level. Also common to H's performance on these items was that when they were first passed, in each case her response only just met the criteria for passing.

Item 90 may have been passed by chance in session 1 at 12 months. At this age level H was particularly interested in noise production and she persistently banged many of the test objects either together or on the table. She had been banging the cube on the side of the cup before she let it drop into the cup. Whether she intended to imitate E's demonstration is questionable but her behaviour fitted the criteria for passing this item.

In session 2 at 12 months, after again persistently banging the cube on the table H did put her hand (and the cube) inside the cup. She would not, however, let go of the cube this time and eventually continued to hit it against the side of the cup. This item was presented a second time in this session. On this occasion H held her hand over the brim of the cup as though she fully intended to put the cube inside it, hesitated and then threw the cube behind her (thereby scoring a fail).

On presentation with the crayon for item 95 H also banged this up and down on the paper. She was however given credit as she had closely attended to E's demonstration and appeared to be watching the marks she was making herself on the page. This was not the case in session 2 when she continued to bang the crayon on the page while looking away in another direction. H responded in a similar way to this item in both sessions administered 3 months later at 15 months. In neither session was there any evidence that she was aware of the marks that she was making with the crayon on the paper.

To pass item 105 the child is required to secure the ring by its string and then to dangle it over the edge of the table in imitation of E's demonstration. In session 1 H secured the ring but could not then coordinate the action necessary to pull the ring off the table. She was credited with this item in session 2 although her success may have been attributable to her having moved by chance away from the table at the right time whilst still holding the string

Items 92: stirs with spoon in imitation; 104: pats whistle doll in imitation

One imitation item which was clearly passed on both occasions at 12 months involved patting the whistle doll (a squeaky toy) in imitation of E's demonstration. This item seemed to excite her at this age level as she was so interested in noise production. At 15 months, however, she did not engage at all in this task; in both sessions she watched E's demonstration and looked away.

H also appeared disinterested in item 92 in the second session administered at this age level. Although keen to demonstrate this activity

2 weeks earlier (and in both sessions at 12 months) she would not attempt to imitate E's demonstration until her mother put the spoon inside the cup for her. She did not therefore receive credit for this item.

Item 106: imitates words

This item was unreliably produced at 18 months and was then dropped from H's protocol 3 months later. At 18 months, during presentation of item 99 (pushes car) H imitated the word 'car'. Her mother reported that H would do this occasionally. She was very quiet in session 2, however, and received no credit for this item or for item 101 (jabbers expressively) which had also been attained in the first session. This lower-level vocalisation item was again passed at 21 months but in neither session was there any evidence that H was attempting to imitate any words. Again H's mother reported that she would do this occasionally - but not on demand.

Object concept items

Items 86: uncovers toy; 96: unwraps cube; 88: picks up cup: secures cube

At 9 months, on presentation with both items 86 and 96, H made no attempt in either case to retrieve the toys once they had been hidden by the tissue used to cover and wrap them. Similarly in item 88, although she did lift the inverted cup and looked at the toy which had been placed underneath it, she made no attempt to secure it. These responses are typical of a child who has not yet developed any understanding of object permanence.

In session 1 at 12 months H at first responded in a similar way to both items 86 and 88. Although she was credited for passing both items there was some delay before she produced the requisite behaviours in both cases. When presented with item 86 she at first seemed to be attending more to the tissue which she stared at for some time before lifting it and securing the toy. Similarly she hesitated before lifting the cup and retrieving the toy in item 88. Although she briefly put the cup to her mouth, her intention seemed clearly to be to reach for the toy. It seemed almost as if she was demonstrating two levels of activity with the cup at the same time. When she put it to her mouth she was exhibiting her recognition of the use of this object. She was also aware of its properties as an occluder, however.

Both items 86 and 88 were passed without hesitation in the second session administered at 12 months.

Three months later, in the first of the two sessions administered at 15 months, H reverted to the level of behaviour seen at 9 months on item 86, simply reaching for and tearing up the tissue. Her response to item 96 (unwraps toy) was similar. No attempt was made in either item to secure the object hidden under/inside the tissue. Rather than attempt to secure the toy which had been hidden under the cup in item 88 H responded by banging the cup on the table and then casting it. When this item was presented a second time she simply hit the toy with the cup, making no attempt to retrieve the toy.

H again preferred to tear the tissue when presented with item 86 in session 2 at 15 months. She did nevertheless clearly pass item 96 for the first time in this session. Item 88 was again failed; H lifted the cup, looked at the toy and then started to wave the cup in the air. She did not reach for the toy until her mother drew her attention to it.

Both items 86 and 88 were clearly passed in session 1 and 2 at 18 months. H did not, however, receive credit for item 96 in either session; on both occasions her response was to tear the tissue. This item was regained at 21 months; in both sessions she very carefully unfolded the tissue wrapped around the cube.

Item 102: uncovers blue box

This item was first passed at 18 months. In previous sessions H had occasionally managed to remove the lid from the box but this had always appeared to be accidental. In the first session administered at 18 months, H very deftly tipped the box with the clear intention of removing the lid and securing the toy which had been put inside it. In session 2, however, she needed several attempts at this item before producing the same successful response. She had twice pushed the box away from her when it was initially presented. She received credit for this item in both sessions administered at 21 months.

Item 131: finds 2 objects

H attained this item for the first time at 21 months but only in the second of the two testing sessions. In this item the child is required to find an object which has been hidden under one of 2 cups. Six trials are presented in total - 3 with a ball and 3 with a toy rabbit. To pass the child has to pick up the correct cup in two out of three trials with each object.

In session 1 H passed all 3 ball trials and the first rabbit trial; in the second rabbit trial, she picked up both cups and in the third, swiped the cup from the table. In session 2 the first two of each of the ball and the rabbit trials were passed successfully. She picked up the wrong cup in the third trial with each object however - although she was still credited for passing this item.

Cube tasks

Items 77: retains 2 of 3 cubes offered; 82: attempts to secure 3 cubes

Having successfully passed items 77 and 82 at 9 and 12 months H dropped both items at 15 months. At this later age level her behaviour did not suggest that she had perhaps surpassed these items but rather that her own criteria for succeeding with the tasks had changed. Credit is given for item 77 if the child retains 2 cubes after being offered a third and for item 82 if he/she then attempts to secure the third cube by "banging at it, reaching with his mouth or by more adaptive scooping, trying to get and hold all 3 at once even though not successful".

At 9 months H passed this item by banging the third cube with one of the two she was holding in either hand in an apparent attempt to secure it. At 12 months her response to the same item had become more sophisticated. In session 1 she not only retained the 2 cubes but also demonstrated her ability to hit them together in the midline. She also succeeded in picking up the third cube when it was presented and then continued to hit the two cubes against the one she was still holding in her other hand. This behaviour was not repeated in session 2 although she was still given credit for item 82 as she reverted to banging the cube with one of the two she was holding, as at 9 months.

At 15 months this response had disappeared. In session 1 she held a cube in each hand and hit them together in the same way as she had done in the first session at 12 months. When presented with the third cube this time she threw the other two cubes onto the table, paused, and then threw them one by one onto the floor. In the second session she seemed unable to retain 2 cubes when a third was put on the table in front of her, dropping one from one hand each time this item was presented.

In both sessions at 15 months H had passed item 100 (puts 3 or more cubes in cup). It could therefore be suggested that because she was able to perform this more complex task with the cubes, she had surpassed items 77 and 82. Two aspects of her performance suggest otherwise. The first is that she had still been keen to demonstrate her ability to hit the two cubes together in the midline, suggesting that she was still interested in this lower level activity with these objects. The second is that she did not reject the cubes until the task became more difficult. Although in session 1 she did not even attempt to secure the third cube she did reveal herself incapable of doing so in session 2 on 3 occasions. In neither session however did she attempt to secure the third cube by hitting it as she had done at 9 months and in the second session at 12 months. In the case of item 82 it may be that H had indeed surpassed this particular response, seeing the more sophisticated response which she had demonstrated in session 1 at 12 months to be more appropriate. If the unreliability of this higher level response at 12 months accurately reflects the level of use of this action of picking up a third object, it would not be surprising if it was not fully consolidated into H's repertoire. It would appear that this may have led not only to this skill being compromised but also to its possible loss.

There may be some connection between H's apparent loss of this skill and her response to a more advanced cube task at 18 and 21 months. At 18 months, although she was very willing to put 9 cubes in the cup (item 114), she was far less cooperative when presented with a task which required a greater degree of motor skill (item 111: builds tower of 2 cubes). Her initial response to E's demonstration and request to imitate this task was to bang the cubes on the table. After a great deal of persuasion she eventually attempted to put one cube on top of the other and succeeded - although with difficulty. She then threw both cubes at her granny who was watching this particular session. Two weeks later she refused outright even to attempt this task and threw the cubes off the table.

Three months later, at 21 months, H was no more cooperative on this tower building task. In session 1 she initially responded to the cubes by picking them up one by one and dropping them onto the table again. She then pushed one off the table. When she did eventually put one on top of another she immediately knocked it off, looked away, made a silly face and dropped the cubes onto the floor. Similarly in session 2, although she did build a tower of 2 cubes, when asked to try to place a third on top of it, she pointed away and then swiped at the cubes. Clearly H was providing herself with little opportunity to improve her performance on this task. Rather than try again, when faced with a failure, or a situation of potential failure, her response to these tasks was simply to reject them.

Summary

It can be seen from the above that H's development in the 12 months in which it was monitored did not show the steady incremental pattern expected by the BSID. In three skill areas in particular - imitation, object concept and cube behaviours - there was repeated evidence both of unreliability of new developmental acquisitions and apparent failure to consolidate these new skills into her behavioral repertoire.

CASE STUDY 2

Subject 12 (N): 9-24 months

Personal details

Sex - M
Karyotype - standard trisomy 21
Health - hearing loss (grommets inserted): otherwise, very good

Test details

Age level	9m	12m	15m	18m	21m	24m
No. sessions	1	2	2	2	2	2

Unreliable items

12m

44 : carries ring to mouth
45 : inspects own hands
50 : manipulates table edge actively
55 : vocalises attitudes
62 : turns head after fallen spoon
65 : smiles at mirror image
70 : picks up cube deftly and directly
91 : looks for contents of box
93 : looks at pictures in book

18m

75 : looks for fallen spoon
81 : cooperates in games
87 : fingers holes in peg board
99 : pushes car along
101 : jabbers expressively

15m

54 : picks up cube
56 : retains 2 cubes
64 : reaches for 2nd cube
67 : sustained inspection of ring
69 : transfers object hand to hand
70 : picks up cube deftly and directly
72 : interest in sound production
75 : looks for fallen spoon
76 : playful response to mirror
78 : manipulates bell : interest in detail
80 : pulls string adaptively: secures ring
83 : rings bell purposively
101 : jabbers expressively

21m

56 : retains 2 cubes
63 : lifts inverted cup
64 : reaches for 2nd cube
66 : bangs in play
67 : sustained inspection of ring
69 : transfers objects hand to hand
72 : interest in sound production
73 : lifts cup with handle
75 : looks for fallen spoon
85 : says dada
86 : uncovers toy
90 : puts cube in cup
92 : stirs with spoon in imitation
93 : looks at pictures in book
97 : repeats performance laughed at
99 : pushes car along
103 : turns pages of book
105 : dangles ring by string
109 : removes pellet from bottle
116 : uses gestures to make wants known

24m

83 : rings bell purposively
87 : fingers holes in peg board
95 : attempts to imitate scribble
98 : holds crayon adaptively
99 : pushes car along
100 : puts 3 cubes in cup
102 : uncovers blue box
104 : pats whistle doll
109 : removes pellet from bottle
122 : attains toy with stick

Unstable items

Items dropped at 12m

77 : retains 2/3 cubes offered
78 : manipulates bell: interest in detail
83 : rings bell purposively

Items dropped at 15m

63 : lifts inverted cup
73 : lifts cup with handle
77 : retains 2/3 cubes offered *
91 : looks for contents of box +

Items dropped at 18m

56 : retains 2 cubes +
64 : reaches for 2nd cube +
66 : bangs in play
67 : sustained inspection of ring +

Items dropped at 21m

77 : retains 2/3 cubes offered *
78 : manipulates bell: interest in detail*

Unstable items *(continued)*

Items dropped at 18m

72 : interest in sound production+

77 : retains 2/3 cubes offered *

78 : manipulates bell
interest in detail +

83 : rings bell purposively +

Items dropped at 21m

83 : rings bell purposively *

87 : fingers holes in
peg board +

Items dropped at 24m

56 : retains 2 cubes +

77 : retains 2/3 cubes offered *

78 : manipulates bell
interest in detail *

85 : says dada +

92 : stirs with spoon in imitation +

96 : unwraps cube

105 : dangles ring by string +

\+ *denotes items which were also unreliable at the preceding age level.*

* *denotes items which were first dropped at a previous age level.*

Subject 8 (N): Analysis of Performance Profiles

This child's profile of scores represents the clearest example found among this group of DS children of a decline in rate of development as measured by the BSID. At 9 months N attained a raw score of 69 which translates into an MDI of 50. Fifteen months later he had only gained 30 raw score points, not all of them reliably attained. Even his optimal score of 99 at 24 months is too low even to be converted into an extrapolated MDI level. The lowest MDI provided by Naglieri (1981) is 28 which would require a raw score of 103.

This child also demonstrated one of the clearest examples of loss of a low level skill. At 24 months of age he could not be given credit for item 56: retains 2 cubes. Despite apparent full engagement in this very simple task he was unable simultaneously to retain a one inch cube in each hand. His mother confirmed that he had lost this ability, one which he had demonstrated at 9 months. He nonetheless passed item 100 (puts 3 or more cubes in the cup) at this age level and made an (unsuccessful) attempt at item 111 (builds tower of 2 cubes).

Before the age of one year N developed the habit of casting virtually every object he was presented with - both in and out of the testing room. His mother had been told at this stage that this was just a phase which he would soon grow out of. This was not unsound advice; many children, handicapped and non-handicapped, do develop this habit which, in normal development, usually disappears around 15 months. N was still casting 12 months later however by which stage the deleterious effects this 'phase' had had on his cognitive and motor development were becoming very apparent. He had become extremely skilled at casting itself and could propel objects at great speed across the room. He had, however, a very rudimentary understanding of the object concept. This is hardly surprising given that he rarely retained any object for long enough to learn anything about its properties. Despite his slow progress on tasks which tested this sort of skill, he did attain a number of social items at each age level which indicated that his level of functioning was not as low at all levels as his BSID scores suggested.

A more qualitative analysis of N's longitudinal BSID performance provides a great deal in the way of explanation for this asynchronous pattern of cognitive and social development. When presented with the cube items at 9 months N was persistent in his attempts both to reach for and grasp the cubes. Despite clear evidence of difficulty he managed to secure 2 cubes and retain one in each hand for a brief period. Three months later, when presented with a single cube, rather than attempt to pick it up his response was to fling it off the table. At this point his mother suggested that he might take the cube if it was handed directly to him. Already at this stage therefore he seemed to have become reluctant to pick up objects by himself. In fact on this occasion he did reach for and pick up a second cube after E had put one in his other hand - although this took several attempts on his part. As N had passed item 84 (listens selectively to familiar words), it was necessary to present a further 6 items beyond this to establish his ceiling level including item 90 (puts cube in cup). Although he did not pass this item his mother reported that he used to put his toys into a box when he was younger but that he no longer seemed to want to do this. In this first 12 months session N failed 8 items which he went on to pass in the second session administered at this age level. Among these was item 70: picks up cube deftly and directly. Despite the initial reluctance and apparent difficulty which he had shown with this task in the first session therefore, it seemed that when he was prepared to try, he could pass it easily. In this second session he was also far more cooperative on item 91 (looks for contents of box), an item in which he had refused to engage in session 1.

Item 91 was among the three items dropped at 15 months. In both 15 month sessions he cast the box across the room rather than attempting to look for its contents. He also on both occasions immediately swiped the cup off the table on presentation with item 63 (lifts inverted cup), whereas in sessions 1 and 2 at 12 months, although he did not secure the toy hidden under it, he deftly picked up the cup using the handle, thereby gaining credit for both items 63 and 73 (lifts cup with handle). He did however manage briefly to retain 2 cubes in between bouts of throwing or dropping them in session 1 at 15 months. In session 2 he failed both this and all other cube items responding by using both hands to scrabble the cubes over the table.

By 18 months N had become almost untestable. In both sessions his response to almost every item which involved any sort of manipulation of an object was simply to swipe the object from the table or to pick it up and throw it across the room. He did, however, pass a number of social items at this age level including item 81 (cooperates in games) and item 89: (responds to verbal request); he also clearly enjoyed the book (items 93/103: looks at pictures in/turns pages of book). In session 1 the only object which he did retain for any length of time was the ring which he pulled towards him to secure (item 80) and dangled over the edge of the table (item 105). His mother mentioned that he had at home a similar toy with a string which he liked to pull. Interestingly, he seemed very unsure of this object when it was presented to him 3 months later, at 21 months. He did tentatively pull and secure it but he would not copy E's demonstration of dangling it over the table. This is particularly noteworthy given that in the first of the sessions at this age N was slightly more willing to manipulate other test materials. Although he did still cast a number of objects he also retained several which he had immediately rejected on presentation in earlier sessions. Two weeks later, however, he returned to casting almost every object across the room - except the ring which he immediately dangled over the table in response to E's demonstration - although he did eventually cast this object also.

During the brief lapse in the casting 'phase' which occurred in session 1 at 21 months N provided some evidence as to the possible cause of this behaviour which had been so prevalent in previous sessions. N's fine motor coordination was very poor. His mother reported that he had recently become interested in putting objects into, and removing objects from containers - but that his success rate on this sort of task was very low. This was clearly apparent from his attempt at item 100 (puts 3 or more cubes in cup). Despite his clear intention to execute this action, having picked up a cube, he had great difficulty in controlling the movement necessary to transfer it successfully to the cup. Rather than a smooth wrist movement he demonstrated more of a throwing action - aiming the cube at the cup. When presented with item 92 (stirs with spoon in imitation), he again seemed to have difficulty with the movement necessary to place the spoon inside the cup. Several presentations were necessary before he succeeded with this item. His coordination was at first so poor that he responded by throwing the spoon in frustration.

As already mentioned, this item, together with almost every other item which involved any manipulation of objects, was failed by default in session 2 at 21 months due to N's return to the casting response. It also disappeared completely at 24 months. In the first session at this age level he did not imitate E's demonstration of stirring with the spoon but instead seemed to be mimicking an eating action. In session 2 he handed both objects back to E without making any attempt to perform the response required to pass this item. As at 21 months, there was a marked difference between his behaviour in sessions 1 and 2 at this age level. Whereas in the first session he was at least willing to attempt several items (items 83: rings bell; 91: looks for contents of box; 107: puts beads in box; 111: builds tower of 2 cubes) these same items were failed two weeks later because he immediately cast the objects on presentation or handed them back to E. Given that he had encountered such difficulty with these items in session 1, his response in session 2 is, in a sense, understandable. Despite his perseverance on items 92 and 111 in session 1, both of which required a high degree of fine motor coordination, he met with little success. In that first 21 months session he also needed a considerable amount of persuasion before he would attempt item 100 (puts 3 or more cubes in cup) and as he was performing this task he pointed away several times in an attempt to divert E's attention away from the task. This item was in fact among the few he attempted in session 2 although in this session, his actions were far less controlled. After initially responding to the cubes by carelessly throwing them across the table he then decided to aim several at the cup and succeeded in placing 2.

Summary

A great deal of fluctuation is evident in this S's test performance, not only in terms of his approach to specific items but also in his general response to being asked to perform actions on objects. At 9 months he was willing to try, despite being rewarded with little success. By 12 months, the effects of persistent failure were beginning to show: he had become reluctant even to attempt tasks which, when he did try, he could occasionally perform successfully. At 15 and 18 months he seemed to have given up altogether and simply cast every object that was presented to him. By 21 months he had become slightly more willing to cooperate, but seemed to have 'off days' when he would return to casting in spite of his growing interest in objects and their properties. At 24 months the casting behaviour was gradually being replaced by a less aggressive, almost more resigned rejection response where he would simply return objects to whomever had presented them to him.

CASE STUDY 3

Subject 18 (S): 36-54months

Personal details

Sex - F
Karyotype - standard trisomy 21
Health - cardiac difficulties; minor hearing loss; slight squint

Test details

Age level	36m	42m	48m	54m
No. sessions	2	2	2	2

Unreliable items

36m

115 : closes round box
117 : shows shoes
122 : attains toy with stick
125 : imitates crayon stroke
126 : follows doll directions
134 : pegs placed in 30 secs

42m

111 : builds tower of 2 cubes
123 : pegs placed in 42 secs
129 : blue board : places 2 round & 2 square blocks
132 : points to 3 pictures
137 : pink board : completes
140 : broken doll : mends approx
142 : blue board : places 6 blocks

48m

117 : shows shoes
120 : pink board:places round block
123 : pegs placed in 42 secs
125 : imitates crayon stroke
129 : blue board: places 2 round & 2 square blocks
132 : points to 3 pictures
135 : differentiates scribble from stroke
143 : builds tower of 6 cubes

54m

120 : pink board : places round block
130 : names 1 picture
132 : points to 3 pictures
135 : differentiates scribble from stroke
142 : blue board : places 6 blocks
147 : imitates crayon strokes: vertical & horizontal
161 : builds tower of 8 cubes

Unstable items

Items dropped at 42m

117 : shows shoes +
125 : imitates crayon stroke +
127 : uses words to make wants known
134 : pegs placed in 30 secs +

Items dropped at 48m

127 : uses words to make wants known *
133 : broken doll : mends marginally
137 : pink board : completes +
140 : broken doll : mends approx +
142 : blue board : places 6 blocks +

Items dropped at 54m

117 : shows shoes +
127 : uses words to make wants known *
133 : broken doll : mends marginally *
140 : broken doll : mends approx *

\+ *denotes items which were also unreliable at the preceding age level*

* *denotes items which were first dropped at a previous age level.*

Subject 18 (S): Analysis of Performance Profiles

It will be recalled from Chapter 5 that this subject produced very different test and retest protocols at both 42 and 54 months (see p). Whereas the scores attained in the first sessions administered at these age levels suggested that no developmental progress had been made, S's optimal score profile revealed that she was in fact acquiring new items over this and indeed every 6 monthly interval. At the same time, however, a large number of items disappeared from her protocol over adjacent pairs of ages which prevented these new acquisitions from making much impression on her scores.

A strong relationship was found to exist between items on which S's performance was found to change over closely spaced intervals and items which dropped out between age levels.

Items dropped between 36 and 42 months

Between 36 and 42 months S dropped item 117 (shows shoes), item 125 (imitates crayon stroke), item 127 (uses words to make wants known) and item 134 (places pegs in 30 seconds). Only item 127 had been reliably produced at 36 months: her performance had changed in a fail-to-pass direction on item 134 and in the opposite direction on items 117 and 125.

Item 117: shows shoes

It will be recalled that item 117 was among those on which many children's performance was found to be both unreliable and unstable. There also appeared to be some correspondence between the emergence of more difficult discrimination skills and the loss of this item. A particularly interesting pattern emerged from S's profile in relation to this specific item.

In the first of the two sessions administered at 36 months, S clearly and willingly pointed to her shoes when asked. Her response in session 2, however, was not sufficiently clear to be credited with a pass; after some delay she simply moved her hand in the direction of her feet but continued to look at E.

When item 117 was presented 6 months later, even this response seemed to have disappeared although it did reappear when the picture cards were presented (item 132: points to 2 pictures). When asked to point to the picture of the shoe S did so and then after some delay pointed under the table at her feet although at the time she was looking in another direction. A similar response was observed two weeks later although on this occasion she refused to point to any of the pictures (she had pointed to 3 in session 1). At 48 months item 117 was passed in both sessions but in session 1 S responded to the picture item by turning the card round and round on the table in an obvious attempt to avoid the task. Item 132 was, however, passed in session 2.

S again rejected the picture cards in the first session administered at 54 months - but this time also did not respond at all to E's request to point to her shoes. This occurred again in session 2 although in this session, when presented with the picture item, she actually named the shoe and then pointed to her own shoe.

One explanation for the apparent disappearance of S's response to item 117 may be that as she became able to point to and discriminate pictures, she became reliant on this pictorial information as an aid in discriminating actual objects. It may be that as she learned to discriminate pictures of objects she was also taught to point to a 'real life' example of whatever was represented in the picture. As she became more competent with this method, however, she lost the ability to directly associate words and objects.

Item 125: imitates crayon stroke

This item is presented after the child has been given an opportunity to scribble spontaneously (item 112). On a new piece of paper E draws a line with the crayon and indicates to the child that he/she is to imitate.

At 36 months S clearly passed item 112 in session 1. She also briefly imitated E's demonstration for item 125 before returning to scribbling. When this item was presented in the second 36 month session, however, she preferred to continue to scribble and would not repeat the crayon stroke produced 2 weeks earlier.

Six months later, when presented with this item in both sessions, S again continued to scribble, leaving E in no doubt that she was simply refusing to cooperate. Her father stated several times during the first session that it was very difficult to know what S could and could not do as she so often refused to perform tasks when asked.

She demonstrated a similar refusal to cooperate in a more advanced task with the crayon and paper at 48 months. In session 1 she was given credit for both items 125 and 135 (differentiates scribble from stroke). The more difficult item could not be presented in session 2, however, as S firstly grabbed the crayon from E's hand (she was already holding one herself) and then when she eventually did return it, refused to allow E to demonstrate a scribble.

A further 6 months later S was again reluctant to cooperate on the crayon and paper items, preferring to continue to scribble than to attend to E's demonstration. After some delay she did pass item 147 (imitates crayon strokes: vertical and horizontal) but this success was not repeated in session 2 when she would not fully attend to the demonstration and was intent on drawing curved lines around the page.

Item 134: pegs placed in 30 seconds

In this item credit is given for the best time achieved in 3 trials at placing 6 pegs in their holes. In the first 2 trials presented in the first session at 36 months S completed the task quite skilfully but took more than 30 seconds to do so in each case. She caused problems for herself in trial 3 by attempting to carry out the task with 2 pegs in each hand. Her eventual response in this case was to cast the pegs. In session 2 S completed the same task within 30 seconds in one of the 3 trials (a credit).

At 42 months she would only complete one trial in each session - in both cases attaining considerably slower times than attained in the second session 6 months previously. When this task was first presented at 48 months, after having placed several pegs, S accidentally knocked the row of pegs which were lying by the side of the peg board and they started to roll across the table. This distracted her and she insisted on putting them back in a neat row before she would continue. Consequently her time was again very slow. In the second and third trials she seemed more intent on removing and replacing the pegs rather than finishing the task. Two weeks later S seemed to have genuine difficulty with this task. After placing 3 pegs she seemed to be unable to aim the remaining 3 at the holes and she therefore pulled out the ones she had already placed. She had to be persuaded to attempt this task a second time completing it in around 45 seconds. She would not give this task a third try.

Items dropped between 42 and 48 months

Between 42 and 48 months S dropped 4 items: items 133/140: mends broken doll marginally/ approximately; item 137: pink board: completes: item 142: blue board: places 6 blocks. Only item 133 had been reliably passed at 42 months.

Items 133/140: broken doll: mends marginally/approximately

Item 133 was also passed on both occasions at 36 months. Because S had placed the doll's head on upside down in session 1 however, she was not also credited with item 140. In fact S attempted to correct her error but encountered difficulty, possibly because the doll's head is so small. In session 2 she placed it on the right way up and was credited for having 'mended' the doll approximately.

At 42 months S passed item 140 in session 1 but in session 2 she could only be given credit for item 133 for her brief attempt at this task which she terminated by handing the doll to E.

At 48 and 54 months she cast the doll every time it was presented to her.

Item 137: pink board: completes

In this task the form board is presented to the child with all 3 shapes (circle, square, triangle) in place. Each shape is then removed and placed opposite its own form on the table between the board and the child.

In session 1 at 36 months S completed this task apparently by using the position cues: rather than matching each shape to its corresponding form she seemed to be using the position of the shapes along the side of the board to guide her. She successfully used this strategy again in session 2 and in the first session administered at 42 months. She could not, however, be given credit for this item in the second 42 month session because, although she put the square and triangular blocks near their forms on the board, she did not fit them correctly.

By 48 months S seemed to have genuine difficulty with this item. There was no evidence of the position strategy used at earlier ages. After a few unsuccessful attempts to place the blocks she eventually lost interest and looked away. In session 2 she immediately cast all 3 blocks without even attempting the task.

Six months later S threatened to reject this task when it was first presented - laughing and pointing away. She did nevertheless settle into an attempt which was not successful. She became very frustrated with the square block which she had placed incorrectly over its form. Rather than adjust its position she hit and pushed at it very hard. She behaved similarly with the triangle which she had also incorrectly placed over its form, responding in the same way to this task when it was presented two weeks later. On this occasion she then pointed away and pushed the board away from her whereas in the previous session she behaved as though she had completed the task despite the two blocks which were not correctly placed.

Item 142: blue board: places 6 blocks

A similar sequence of unsuccessful attempts and rejections was observed in S's response to this second form board task which involves placing 5 round and 4 square blocks. Whenever she encountered difficulty she simply either cast the blocks or pushed the board away from her. Her most successful attempt at this task was seen in session 2 at 42 months when she correctly placed 6 of the 9 blocks. On this occasion she initially appeared to know exactly where each block should be placed. Having made one mistake, however, she seemed to lose confidence in her ability to complete the task and after making several more unsuccessful attempts, rejected it. At 42 months, rather than adjusting the position of incorrectly placed blocks, she attempted to fit them by pushing them very hard against the board, as in item 137.

Interestingly, at 54 months, S seemed to be attempting to apply to this task the strategy which she had used at earlier ages on the pink form board task. Unlike that task in which all the blocks are placed in front of the child, examiners are instructed in this item to hand each block to the child individually. Having been handed one block, S refused to continue with the task until the remaining 8 blocks were placed alongside the board (although not in line with their matching hole). This arrangement did not, however, seem to be for the purpose of visually matching the blocks and their forms - her attempts to push round blocks into square holes would seem to indicate that she did not understand the concept of shape. Instead it appeared that S was working on the same principles which she had earlier applied to the 3-shape form board task.

Summary

From this child's longitudinal profile it is possible to see a clear link between unreliable and unstable performance. A large number of the items on which her performance was seen to change over closely-spaced intervals were also those which later dropped out. The repeated evidence of unsuccessful attempts at many of these tasks strongly suggests that the link between early unreliability and subsequent instability may be a causal one with failure to reliably produce skills when first acquired leading in turn to a failure to consolidate these skills adequately into the behavioral repertoire.

OVERVIEW

The aim of the analyses presented in this chapter of DS longitudinal performance on the BSID was to focus on the possible relationship between unreliable performance on specific BSID items and the apparent later loss of these same items from the DS child's behavioural repertoire.

As was the case in the smaller-scale cross-sectional study reported in Chapter 4, DS test-retest reliability was found to be significantly lower than that reported for non-handicapped children. This variability of performance was found, to a large degree, to be attributable to Ss' failure to engage in tasks in one or other of the two sessions. In fact the prevalence of such failures by default may have meant that even the low level of reliability found is an overestimate. Given that unreliability does seem to be characteristic of the performance of DS children, both in and out of the testing room, there is little reason to assume that even two short testing sessions will necessarily demonstrate its full extent. In this study failures by default were often found to have occurred on *both* occasions of testing, leading to 'reliable' performance yet clearly producing a measure which did not reflect the extent to which children were not performing to full competence.

There is equally little reason to assume that repeated failure to reproduce items which had been passed at earlier age levels necessarily indicates that Ss have genuinely lost these skills from their behavioural repertoires. As suggested in the introduction item loss may simply be explained in terms of failure on the part of the test materials - which were selected for use with much younger non-handicapped children - to interest the older handicapped child. Alternatively, as Bayley suggested, the level of the item itself may be inappropriately low for the child who has advanced onto more complex activities with the objects presented.

Neither of the above explanations can, however, account for the finding that failure to engage was found to account for only 16% of instances of item loss. The result of this particular analysis contrasts markedly with that found for the analysis of unreliable performance over closely-spaced testing sessions in which failure to engage was recorded in 64% of cases. Although, in itself this finding cannot necessarily be interpreted to indicate that skills were genuinely lost, evidence emerged from subsequent analyses to support the suggestion that they were indeed lost.

More qualitative analyses of failures to engage in fact indicated that Ss were not avoiding specific items because the test materials themselves were inappropriate. Responses often seemed to vary more in relation to the tasks than to the objects used in certain items. Failure to engage was far less prevalent, for example, on item 102 (uncovers blue box) than it was on item 107 (puts beads in box). Almost 60% of children who were presented with the latter item either refused to attempt the task of placing 6 small beads into a small hole in the lid of the box, or 'opted out' of the task before completing it. Far fewer children (13%) rejected this same object after watching E hide a small toy inside it. In fact many of the older children in this sample enjoyed playing with the blue box when it was presented for the discrimination item (144/ 152: discriminates 2/3: cup, plate, box). Items 144/152 were nevertheless among those most frequently failed by default; a common response to the objects was to play 'tea sets' rather than to cooperate in the discrimination task. There was a similar pattern with the jointed doll. Item 126 involved following E's instructions to place the jointed doll on a chair, give it a drink and wipe its nose. This item was a particular favourite and yet this same doll became especially unpopular when the task required pointing to parts of its face and body.

From the above examples, it would appear that something other than item-unsuitability was responsible for the avoidance of certain items by DS Ss. A more likely explanation is that by failing to engage in such items Ss were avoiding the possibility of failure. In doing so it is possible that they were depriving themselves of a vital element in the process of consolidation (see below).

The absence of evidence to suggest that children had surpassed these items also supports the argument that they may genuinely have been lost. As a large number of cases of item loss involved the sorts of low-level cube skills which Bayley herself had identified to be those which older children are likely to have surpassed, these items were given particular attention in the analyses carried out here. In a substantial proportion of cases in which such items did drop out there was little evidence to suggest that this had occurred because children had advanced onto more complex activities with the same task objects. Even in cases in which a low-level cube item was dropped and replaced by a more complex item, almost without exception the higher level item was not produced reliably. This finding is very much at variance with the implicit suggestion that by surpassing low-level items children are showing themselves to be motivated to demonstrate their newly acquired higher level of skill. Failure to reproduce newly acquired skills reliably, by contrast, would seem to indicate that the DS child is not as motivated - which in turn may imply that s/he has encountered some difficulty in acquiring these abilities in the first instance.

Moreover, there is reason to suggest that even in cases where lower level items are replaced by items of apparently higher levels of complexity, this does not necessarily indicate that the absent item is still within the child's repertoire of skills. As Garwood (1982) stresses, the order in which items appear in the BSID is not based on any theoretical developmental sequence but is instead based on relative difficulty levels. Items are arranged on the basis of their ability to reflect an increase in the percentage of passes with increasing age. This sequence, which was determined by the performance of the BSID standardisation sample does not therefore imply that the skills tested should be acquired in an invariant order. Bayley herself stated that the sequential order only holds when the items are sufficiently far apart on the difficulty scale to constitute manifestations of different levels of ability. In the absence of any guidelines as to how to determine whether items are sufficiently far apart, the issue of surpassed items becomes somewhat ambiguous. Although, as stated above, items testing cube behaviours were specifically identified as the sorts of items which children are likely to surpass, no explanation is offered for the overlap between the range of age placements covered by lower and higher level cube items. Item 82 (attempts to secure 2 of 3 cubes offered), for example, has an age placement range of 5 - 14 months which overlaps quite considerably with the age placement range of 10 -19 months provided for item 111 (builds tower of 2 cubes). This degree of overlap would seem to indicate that the two items are not sufficiently far apart on the difficulty scale to imply any sequential order in acquisition - which in turn implies that ability to pass the 'lower' item is not a necessary prerequisite for passing the 'higher' one. In addition the overlap between the age placement ranges indicates that some children in the standardisation sample may indeed have attained these two skills in 'reverse' order.

There is no reason to assume, therefore, that by passing item 111 the child has necessarily demonstrated that the skills required to pass item 82 are within his/her repertoire. If the 'lower' level item is not a pre-requisite for the 'more complex' item 111, it may well be the case that a child has not yet acquired or may even have lost the ability to pass item 82, but has gained the ability to build a tower of cubes. This suggests that the issue of surpassed items should be viewed with extreme caution by examiners testing groups of children suspected to have consolidation difficulties. Rather than having advanced onto a higher level of complexity of skill it may simply be the case that the DS child acquires one skill, only to lose another, related skill

If it *is* the case that skills are dropping out and are being replaced by other, similar skills this could indicate that the DS child has not seen how the two skills interrelate. Whether or not they are acquired in any invariant order it is likely that two similar skills should have some carry over of understanding from one to the other. This point may best be illustrated by another example. At 15 months, H - one of the Ss followed as an individual case study - dropped two object concept items (item 86:

uncovers toy; age placement 6 - 12 months; item 88: picks up cup: secures cube; age placement 6 - 14 months) but gained another item testing the same kind of skill (item 96: unwraps cube; age placement 8 - 17 months). Because, as with the cube items discussed above, the age placement ranges overlap, there is no reason to assume that these items should be acquired in the order in which they appear in the test. For the same reason, however, simply by demonstrating herself capable of passing item 96, H cannot be assumed to have necessarily surpassed the two 'lower level' items. It might, nevertheless, be argued that the child is likely to apply this skill to a task which requires removing an object from inside another object, a tissue, for example, having initially acquired the simple skill involved in retrieving an object hidden under the tissue.

This did not seem to have occurred in H's case. Instead her performance indicated that she had lost the 'uncovering' skills but gained the 'unwrapping' one - suggesting that she did not understand the way in which the two were related. Rather than having acquired an understanding of the object concept, what may instead have been acquired was a set of strategies for carrying out each separate task. Such an explanation lends support to Morss' suggestion that the DS child learns in an incomplete way - that success reflects more of an ability to perform the required actions than an understanding of the concepts underlying these actions.

A similar explanation may apply to the relationship which emerged between unreliable and unstable performance. In a substantial proportion of instances of item loss it was found that the same items had been passed on only one occasion at the preceding age level. Not every item on which performance was seen to change over closely-spaced intervals was subsequently lost and not every case of item loss was preceded by an unreliable performance at an earlier age but a high degree of correspondence between these two performance factors was identified. Given that unreliability does seem to be a characteristic of DS children's approach to many cognitive tasks both in and out of the testing room it may well be that by failing to regularly rehearse and explore newly acquired skills the DS child does not adequately complete the learning process necessary for the consolidation of newly acquired skills. If unreliability and instability are causally linked in this way, this strongly implies that attention should be focussed on unreliability as a problem in itself; on its possible causes and on the ways in which it may prevent consolidation of new learning.

Findings presented earlier in this thesis have indicated that the tendency to avoid cognitive tasks may be related to an imbalance in the success/failure ratio experienced during skill acquisition. Unlike the normally developing child who is able to maintain a sufficient level of success to provide the incentive to continue to learn, it has been proposed that the DS child, by nature of the handicapping condition, too often experiences failure and thereby loses the motivation to persevere. In this study, however there was evidence from the qualitative analyses of DS performance, that the origin of this problem may not always be cognitive. The performance of the second child followed as a case study - N - can be used to illustrate this point. N was often unable to perform a given task not, it appeared, because he did not have the cognitive skills to pass that task but because of his motor disabilities. He clearly demonstrated his intention to learn about the properties of objects as containers - his mother's report confirmed this intention - but his attempts to explore this new interest were hindered by his inability to coordinate his actions. N's response to failures of this sort was to reject the task.

Given that erring is believed to be as essential a part of the learning process as success, it is possible that by constantly avoiding failure N was depriving himself of a vital source of information about the properties of containers - his own errors. By attending only to his successes he was able to learn only what can be done with containers and nothing about the properties that they do *not* possess. In this way N was preventing completion of the accommodation - assimilation sequence which Piaget has proposed is necessary before a child can achieve a full understanding of why it is that a successful action is successful. The majority of N's failures could not be termed as 'errors' in the Piagetian sense - they were brought about by his motor difficulties. Nonetheless there seems no reason to assume that his response to the sorts of errors which are encountered in normal cognitive development - errors which normally developing children capitalise on and learn from - would be any different from those which were due to his inability to coordinate his actions, i.e to reject them.

Many other Ss demonstrated similar levels of difficulty with the motor components of specific tasks. Item 107 (puts beads in box), it will be recalled, was avoided in almost 60% of cases in which it was presented. To be credited for passing this task the child is required to place a maximum of 6 (out of 8) 1 centimetre beads into a hole (approximately 2 cm. in diameter) in the lid of a box. By comparison item 114 (puts 9 cubes in cup), a very similar task was only avoided in 7 out of 38 cases (13%). Although this particular item did present difficulties for several children (N being a clear example), it was rejected less often - possibly because the objects used in this task were easier to handle than those in item 107.

This is not to imply that the cognitive deficit in DS can be solely attributed to difficulties encountered as a result of the motor deficit. DS is considered primarily to be a mentally handicapping condition with associated motor difficulties. The significance of the motor deficit is in its role in bringing about errors which have little contribution to make to the overall process of development. Instead of making a positive contribution to learning by clarifying what has still to be learnt, errors merely add to the child's awareness of his/her own difficulties in learning situations. The response to this experience would seem to be simply to avoid entering such situations.

It will be recalled that Cunningham (1979) reached a similar conclusion from his study of early reaching in DS babies. He explained the frequent absence of the visually directed stage in reaching in early infancy in terms of a mismatch between intention and ability. Poor success rates experienced during attempts to execute the intended action, it was proposed, resulted in the eventual extinction of the reaching response. The results of the present study suggest that the mismatch between intention and ability may continue throughout early development in DS children - with the eventual outcome being failure to consolidate skills and a subsequently poor level of understanding of the interrelationships between differing levels of similar skills.

If this is the case, it clearly implies that attention should be directed not only towards encouraging the development of new skills but also at ensuring that the learning process is allowed to run its full course - that newly acquired skills are fully learned and fully consolidated. The next chapter will return to the starting point of this thesis, errorless learning, in an attempt to determine whether by artificially increasing the success/failure ratio experienced by DS children in learning, it is possible to prevent the avoidance response and its subsequent adverse effects on the consolidation process.

CHAPTER 7

INTRODUCTION

It will be recalled that in the first study presented in this thesis (study 1, Chapter 2) the efficacy of two differing strategies for teaching discrimination were compared: errorless and trial-and-error learning. In that study children in the DS group were observed to respond positively to the errorless learning strategy, attaining comparatively better scores, both in training and in post-tests, than in a similar trial-and-error task. This is perhaps unsurprising; the recent popularity of the technique can be taken as an indication of its effectiveness in teaching skills which children with mental handicap have had difficulty learning by trial-and-error. However, it will be recalled that a more interesting pattern of results emerged from analysis of the effects of order of presentation of the two strategies: initial experience of errorless learning had an enhancing effect on DS performance on a subsequently presented trial-and-error task, with children presented with the tasks in this order producing consistently better scores both on the errorless and trial-and-error tasks than children experiencing these two tasks in reverse order. There was little evidence of such an effect being present in the scores of the non-handicapped group indicating a difference between the two populations in the extent to which performance is affected by differential levels of success and failure.

The studies to be presented in this chapter aim to extend these earlier findings; firstly, by investigating whether any gains demonstrated from the use of an errorless procedure have any real learning value in terms of a/ the stability of what has been learned and b/ the transferability of that learning to other related tasks, and secondly, by comparing the response to errorless learning of two mentally handicapped groups of differing aetiologies, a DS group, and a non-DS group. As the results from study 1 indicated that the errorless learning strategy had no particular advantages for use with the non-handicapped group there seemed little point in including another control group of normal children. It was therefore decided to use a control group of non-DS mentally handicapped subjects. (The reasons for introducing this second group of mentally handicapped subjects will be discussed in full below).

This thesis has placed particular emphasis on the unreliability of the performance of children with DS in assessment situations. In the studies presented in Chapters 3-6 children were frequently observed to withhold demonstration of optimal performance on specific test items. This, it has been suggested, may be due to the increased experience of failure encountered during the early stages of learning, with DS children developing a tendency to avoid situations in which they may encounter further experience of failure. It has also been suggested that poor engagement may be a characteristic of learning in general in DS, affecting both the acquisition and consolidation stages of learning. Teaching strategies such as errorless learning could therefore play an important role in preventing this tendency to avoid situations in which failure is likely thereby overcoming poor engagement both during skill acquisition and once specific skills have been acquired. Artificial enhancement of success/failure ratios could prevent avoidance becoming the routine response to difficult learning situations. A possible outcome of this more 'balanced' experience of success/failure may be that children might become less reluctant to reproduce the skills they have learned, thereby increasing the probability that these skills will be fully consolidated.

Although results from the errorless learning study presented in Chapter 2 were encouraging, they were not sufficient as proof of the intrinsic value of a success-only strategy as a method of enhancing learning ability. Throughout this thesis a great deal of emphasis has been placed on the role played by erring in the process of learning. Support for this viewpoint was offered by Morss (1979) who suggested that learning in DS children often seemed to be incomplete and that this may be attributable to their tendency to avoid learning situations in which failure is likely to be encountered. This failure to appreciate the significance of errors means it may not be possible for these children to fully understand success when it does occur. Teaching children to learn in the absence of error would not therefore seem an ideal solution to this particular problem.

In a similar vein, however, if, due to an exaggerated experience of failure, DS children are reluctant to demonstrate particular skills during the initial stages of their acquisition, it is equally unlikely that any successes achieved will be fully capitalised on. Perhaps in the case of these children it is more important to ensure success, whether or not it is initially fully understood, in an attempt to increase the reliability of production of specific skills. Having confidently 'mastered' these skills through repeated success, the child may be in a better position to deal with the errors necessary for the completion of the learning process. Indeed this was implied by the superior trial-and-error results attained by children who had had prior experience of errorless learning in study 1.

Given the recent popularity of the errorless learning strategy, it seems surprising that so few studies have attempted to assess its long term or its carry over effects. Of the small number of case studies which have reported positive post-training results, there appears to have been little or no mention of the stability of achievements made with this technique. The first of the two experiments to be presented here addresses this issue by investigating whether skills learned in an errorless way will be reliably demonstrated in re-tests up to six weeks following initial training. The second will investigate whether children can transfer the discrimination skills learned with the errorless procedure to a second, errorless task with any degree of saving.

Both studies will compare the efficacy of the errorless technique in two distinct populations of mentally handicapped children. Until now this thesis has focussed on a single aetiological form of mental handicap. Because DS is identifiable at birth, this section of the mentally handicapped population is particularly suitable for a developmental investigation of the relationship between motivation and cognitive development. Although the mentally handicapped in general have frequently been observed to demonstrate poor motivation, further research would be necessary to determine whether the competence/performance patterns found in the studies here also exist in mentally handicapped groups of other aetiologies. Comparison of the response of DS and non-DS mentally handicapped children to a strategy designed to manipulate motivational variables may, however, provide some indication of any similarities or differences in the extent to which non-cognitive factors may be affecting cognitive performance in other mentally handicapped groups. If performance factors prove to be similar this would suggest that findings from DS research studies may well be generalisable to other sections of the mentally handicapped population.

In study 1 errorless procedures were used to teach discriminations both of a common shape and of a previously unseen nonsense figure. For three reasons it was decided in the present study to use nonsense stimuli only. Firstly, this avoided any possible confounding effects on test performance of differential levels of prior learning experience (see Chapter 2 p.40); unlike common concepts such as shape or colour, it could be guaranteed that all subjects were equally unfamiliar with the nonsense stimuli. Secondly, it allowed investigation of the effectiveness of the procedure itself, in isolation from any possible effect of its interaction with experiential factors; no child could have had any previous learning history - either positive or negative - with the discrimination under study. The use of stimuli with which all subjects were equally unfamiliar had the third advantage of allowing a wide age range of subjects to be tested. It will be recalled that in study 1 subjects had been selected for inclusion in the experiment if they were able to demonstrate some knowledge of shapes but were unable as yet to discriminate the particular shape, a rectangle, used in training. Such a selection process could obviously not distinguish between subjects who were performing to full competence in the selection pre-test and truly did not yet know how to discriminate a rectangle, and any who were in fact able to discriminate the shape but, for whatever reason, would not demonstrate this ability. As findings from study 1 had indicated that older subjects were more likely to underperform, the use of nonsense stimuli has the advantage that it places no limitations on the range of subjects' mental and chronological ages, thereby allowing investigation of any possible age effects.

In summary the aims of the present study were as follows:

1. to attempt to replicate the success of the errorless learning technique demonstrated in study 1 with a wider age range of subjects.

2. to investigate any differences between DS and non-DS mentally handicapped children in response to errorless learning.

3. to investigate the long-term stability of achievements made with errorless learning in the two handicap groups.

STUDY 6a - THE STABILITY OF ERRORLESS LEARNING

METHOD

Subjects

Eight DS and eight non-DS mentally handicapped children took part in this study. Both groups of children were selected from three Edinburgh special schools. The DS group consisted of 4 females and 4 males; age range 6 - 16 years (mean age 11.25 years). The age range of the 5 female and 3 male subjects in the non-DS group was 6 - 17 years (mean age 11.5 years).

It will be recalled that, in study 1, direct mental age matching of handicapped and non-handicapped children was avoided on theoretical grounds and subjects matched by means of a task-related selection test. Because nonsense syllables were being used this was not possible in the present experiment. The two subject groups were matched therefore on the basis of teachers' ratings. Teachers were informed of the aims of the study and asked to select non-DS mentally handicapped children who were similar in general ability levels to the children with DS selected for study.

It was felt prudent also to obtain a second measure of overall ability level in the event of two teacher-matched children performing very differently in the discrimination tasks. The accuracy of teacher matches was therefore assessed by determining the mental age levels attained by Ss on the Kaufman Assessment Battery for Children (K-ABC) and comparing these across matches. The mental age matching of two groups of mentally handicapped children, although by no means ideal, is perhaps less vulnerable to criticism than MA matches of mentally handicapped children to non-handicapped children who inevitably are very much younger. The differences expected between the two groups of handicapped children in terms of individual learning histories and motivational factors would be less than between a handicapped and a non-handicapped group. To some extent this might compensate for the fact that their abilities are being measured with reference to those of a 'normal' population i.e. by a psychometric assessment test.

Procedure

Procedure was identical to that used in the errorless nonsense figure discrimination task presented in study 1 (see p) except that pre-tests were included in the present study. This was done for two reasons: firstly, it enabled evaluation of within-session effects of errorless training; secondly, it provided a measure of the long-term stability of skills learned in training through comparison of pre- test scores obtained in sessions 2, 3 and 4 with session 1 post-test scores.

Trial sets each consisted of one target stimulus - a 'wug' - together with two alternative nonsense figures. In training the alternative stimuli were 'faded in' on the basis of size.

The experiment took place over a 6 week period. In the first session, after being presented with the pre-test, all subjects were trained by errorless learning procedures to discriminate the 'wug' from two alternative nonsense stimuli. They were given the post-test immediately following training. Re-tests were administered a week and a fortnight following initial training sessions. Subjects were retrained in the discrimination if they did not achieve 100% on the pre-test at the beginning of each of these re-testing sessions. Pressures of time-tabling meant that DS subjects only could be presented with a further re-test three weeks after the third session (i.e. 6 weeks after the initial training session).

RESULTS

All responses in the pre-tests, training and post-tests were used in the analysis. As in the earlier errorless learning study, pre- and post-test scores were expressed as the number of correct responses made and training scores were expressed as percentages (correct responses/total responses).[1]

Subject matching

Performance in the initial session was investigated for differences that may have indicated any inaccuracies in matching across the DS and non-DS groups. No significant differences were found between pre- or post-test raw scores attained by the two groups ($t = 0.406$, df 14, NS; $t = 0.457$, df 14, NS respectively - see Table 7:1). A further comparison of pre-/post-test differences in individual subjects did not reveal any trend in favour of either group ($t = 0.114$, NS) indicating that for the purposes of this study, subject matches were satisfactory. Comparison of mental age scores attained on the K-ABC also revealed no significant differences between groups, adding further support for the validity of teachers' original ratings.

---

[1]Thanks to Rachelle Walker for her help in the collection of some of the data reported in this chapter (see also Walker 1989).

Table 7:1

'Wug' Discrimination: Effects of Training on Subsequent Performance Pre- and Post Test Scores

DS Subjects

Session:	1		2		3		4
	Pre	Post	Pre	Post	Pre	Post	Pre
1	1	7	7	-	7	-	7
2	0	6	4	5	2	7	7
3	3	7	7	-	7	-	7
4	3	6	7	-	7	-	7
5	4	7	7	-	7	-	7
6	1	7	7	-	7	-	7
7	3	7	7	-	7	-	2
8	3	2	4	5	7	-	7

Non-DS Subjects

Session	1		2		3	
	Pre	Post	Pre	Post	Pre	Post
1	3	7	2	7	7	-
2	1	7	7	-	7	-
3	3	7	3	6	7	-
4	2	5	2	7	7	-
5	0	7	7	-	7	-
6	6	7	7	-	7	-
7	2	7	7	-	7	-
8	4	6	7	-	7	-

Session 1

Training scores

Table 7:2 shows the percentage of correct responses during training trials for both DS and non-DS groups over 3 sessions. Although both groups responded positively to the errorless strategy in the first session, t-test comparison of training scores attained in this session revealed a significant difference in favour of the DS group ($t = 1.974$, df 7, $p < 0.05$). It can also be seen from Table 7:2 that 3 subjects from each group required retraining in the second session and 1 DS subject had to be trained a third time. In all cases, however, training scores were seen to improve in these subsequent sessions, demonstrating at least some level of carry-over effect.

Table 7:2

Percentage of Correct Response During 'Wug' Training Trials

Session:		**1**	**2**	**3**
DS Subjects	1	100	-	-
	2	80	93	100
	3	100	-	-
	4	100	100	-
	5	100	-	-
	6	100	-	-
	7	100	-	-
	8	80	90	-
Session:		**1**	**2**	**3**
Non-DS S's	1	74	95	-
	2	100	-	-
	3	67	73	-
	4	85	100	-
	5	74	-	-
	6	100	-	-
	7	100	-	-
	8	67	-	-

Pre- and post-test scores

Table 7:1 shows the effect of training on post-test performance and on pre-test performance in subsequent sessions. Improvement in performance in session 1 was calculated by comparing pre- and post-test scores. This difference was found to be highly significant both when scores from the two groups were combined and when treated separately (combined groups $t = 7.456$, df 15, $p < 0.0005$; DS group $t = 4.651$, df 7,

$p < 0.005$; non-DS group $t=5.657$, df 7, $p < 0.0005$). Consistent with findings from the errorless learning study in Chapter 2, this initial comparison demonstrates the immediate carry-over effects of a single errorless training session on post-test performance

Pre-test scores attained in the two subsequent sessions provided a measure of the extent to which the discrimination had been retained over the weekly and fortnightly intervals. No significant differences were found between scores achieved on the pre-test in session 2 and those in the post-test in session 1 either when the two groups' scores were combined or were taken separately (combined groups: $t = 1.3$, df 15, NS; DS group: $t = 1.7$, df 7 NS; non-DS group: $t = 0.3$, df 7, NS), indicating that the discrimination had been retained between the first and second sessions. Comparison of the results obtained by the two groups, however, suggested that the DS group may have benefited to a greater extent from the initial errorless training session: from Table 7:1 it can be seen that only one DS subject actually produced a lower score in the second of the two sessions, two other subjects showing score improvements; scores from three non-DS subjects by comparison were reduced in session 2, with only one subject attaining a small improvement in performance. Evidence of retention of the discrimination was, however, clearly demonstrated for both groups in the third session; again, no significant differences were found between pre-test scores in this session and post-test scores from session 1, either overall or for the separate groups. In this third session all but one DS subject attained 100% on the pre-test. After retraining this single subject did succeed in correctly discriminating the wug in all 7 post-test trials and was able to repeat this high level of performance in the final pre-test administered 3 weeks later. As can be seen in Table 2, 7 out of the group of 8 DS subjects had fully retained the discrimination over this second 3 week interval.

DISCUSSION

This study aimed to replicate and to extend the study presented in Chapter 2. In that study errorless learning was shown to be an effective method of teaching discrimination skills to DS children. Here the errorless procedure was investigated for its effectiveness as a means of increasing the reliability of these skills once taught. As expected, in both the DS and non-DS groups there was a significant carry-over effect from training to post-test performance in the first session. In almost 70% of cases a single training session proved sufficient for subjects to select the correct target stimulus in all 7 post-test trials administered immediately following training. With one exception, the remaining subjects, whilst unable to demonstrate full mastery of the discrimination were able to show a marked improvement on pre-test scores after exposure to training. Findings from study 1 were therefore successfully replicated, with errorless learning once again shown to be an effective method for teaching discrimination skills within a single teaching session.

Results also clearly showed that the majority of subjects were still able to discriminate the target stimulus successfully in pre-tests given 1 week then 2 weeks later. Success in these latter two sessions is particularly encouraging given that pre-tests contained only 7 trials. Children in neither group apparently required any prompting to recognise the task and then go on to perform at a similarly high level to that demonstrated after initial training. The fact that 7 out of 8 DS children demonstrated retention of the discrimination when presented with a third re-test 6 weeks after the first training session is particularly encouraging, especially given that in most cases only one training session had been required. Teachers had indicated in preliminary discussions that children in both groups often required repeated teaching sessions when learning similar skills using conventional methods. These results suggest that for both DS and non-DS mentally handicapped groups, skills taught in this way can not only be learned comparatively quickly but are also retained.

A clear difference did, however, emerge between the two groups in training. In the first session DS children required fewer trials to acquire the discrimination than children in the non-DS group. This result may be attributable to the superior visual discriminatory abilities observed in DS subjects in comparison to subjects with other forms of mental handicap (Gordon 1944; O'Connor and Hermelin 1961; Bilovsky and Share 1968; Snart, O'Grady and Das 1982). Given however that this difference was not significantly reflected in any subsequent session, nor in the amount of improvement made between pre- and post-test scores in the initial session, it would appear that these differences demonstrate a more positive response within the DS group to errorless training per se, rather than any inherent advantage in terms of discriminatory abilities. This finding may reflect a stronger relationship between motivational factors and the demonstration of competence in performance in children with DS than in non-DS children. There were several clear examples of DS subjects at all ages 'switching on' social skills in non-training trials - in a way very similar to that so frequently observed in the DS children tested on the BSID - to avoid the task in hand. Little evidence of the use of similar avoidance tactics was observed in the performance of non-DS subjects.

Despite DS children's use of this type of behaviour during testing, it proved to have little effect on the stability of their performance. Although there were subjects in both groups whose scores had dropped between sessions 1 and 2, for the DS group this was only found to occur when the discrimination had not been fully mastered in the initial session. All but one DS subject, having attained 100% on any pre- or post-test, were able to reproduce this score in subsequent tests. This was not always the case with children in the non-DS group where a few quite severe score losses indicated that the discrimination, apparently mastered in the first session, had not been retained. This trend in favour of greater stability of performance among the DS subjects occurred in spite of the tendency among subjects in this group to divert attention from the tasks. Although still in evidence within this learning situation therefore, motivational factors seemed to have a far weaker effect on overall performance than was observed in the earlier studies.

Although encouraging, the fact that subjects responded so well to the errorless procedure does not in itself necessarily mean that this procedure is effective in terms of learning per se. In itself all-correct within-task performance on a discrimination task after errorless training does not necessarily imply that a discrimination skill as such has been acquired. The true value of discrimination learning lies in the ability to transfer that skill to other tasks..

It will be recalled that study 1 demonstrated that prior experience of errorless training in a shape discrimination task had an enhancing effect on DS childrens' performance on a subsequently presented trial-and-error task requiring use of the same discrimination skill. The major purpose of that particular study was not to teach children how to discriminate shapes, however; children were selected for inclusion in the study on the basis that they were already able to identify some common shapes but not the shapes used in the experiment. In fact it will be recalled that the lower level of success attained by subjects on the shape task, when compared to the nonsense figure task, was interpreted as resulting from the very fact that children had had prior experience of shape discrimination learning. The major focus of that study was on the influence of non-cognitive factors on children's performance on discrimination tasks, rather than on discrimination as a cognitive ability per se. These results therefore could also tell very little about the efficacy of errorless learning as a device for teaching discrimination skills. The purpose of the next study was therefore to investigate what it is that is learned with the errorless procedure and whether that learning can be successfully transferred to a second errorless task - one which involves a different discrimination but requires the same basic discrimination skills.

STUDY 6b: THE TRANSFERABILITY OF ERRORLESS LEARNING

METHOD

Subjects

Seven of the 8 DS subjects from study 6a took part in this study. Due to medical appointments DS subject 8 was unable to participate. He was therefore replaced by a 9 year old boy with DS. This lowered the mean age of the group to 10.3 years.

Procedure

Seven sets of cards similar in format to those used in study 6a were used, each including one 'nim' (the target stimulus) together with two alternative nonsense stimuli. One trial set in each of the pre- and post-tests included a 'wug' (the target stimulus from study 6a) as an alternative.

Subjects were tested once weekly for three weeks. As in study 6a, all subjects were trained in the new discrimination in session 1 (after administration of the pre-test, and followed by the post-test). Re-training was given only where necessary in the two subsequent sessions (i.e. if pre-test scores were not 100% correct).

RESULTS

Training scores

Training scores are set out in Table 7:3. It can be seen that most subjects responded well to the change in target stimulus and were able to produce training scores in session 1 of an equivalent level to those achieved in the first training session in study 6a. No significant differences were found between scores attained in the initial training sessions in studies 6a and 6b (t = 0.894, NS). This result suggests that there was little transfer of skill between the two discrimination tasks; had a skill been transferred some savings in the number of training trials required for mastery of the second discrimination would have been expected.

Table 7:3

Percentage of Correct Responses During 'Nim' Training Trials

Session		**1**	**2**	**3**
Subject	1	94	73	89
	2	94	Abandoned	74
	3	Abandoned	74	94
	4	100	100	100
	5	100	-	-
	6	100	-	-
	7	100	100	100
	8	100	100	-

In addition, although in many cases training scores in the second and third sessions were consistently high, it can be seen that many subjects did require re-training. Six subjects received training a second time and five a third time whereas in experiment 1 only three children had required a second, and one a third training session before being able to achieve 100% pre-test scores in all subsequent sessions. It will also be noted from Table 7:3 that 2 training sessions had to be abandoned. This will be discussed below.

Pre- and post-test scores

Table 7:4 shows the effect of training on pre-and post-test performance. Direct comparison of pre- and post-test scores attained in the first session yielded no significant differences ($t = 1.843$, df (7), NS). Whereas in study 6a performance had significantly improved after a single training session, this effect was not found when the target stimulus was changed for the present study. Pre-post test improvement in the first session of study 6a was significantly higher than that achieved on this second discrimination task ($t = 3.308$, df (7), $p < 0.01$).

Table 7:4

'Nim' Discrimination: Effects of Training on subsequent Performance Pre- and Post-Test Scores

Session		**1**		**2**		**3**	
		Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**
Subject	1	4	6	3	5	6	7
	2	0	1*	0	-	0	4
	3	2	-	1	3	5	6*
	4	0	3*	1	6*	4*	7
	5	4	7	6*	6*	7	-
	6	7	7	7	-	7	-
	7	3	5	3	5	5	6*
	8	7	7	5*	6*	7	-

* Denotes sessions in which 'wug' was incorrectly selected

Due to the absence of any immediate carry-over effect onto post-test performance in this initial session, it was not possible to repeat the investigation of longer-term effects carried out in study 6a. Instead statistical comparisons were restricted to those made directly between scores achieved in experiments 1 and 2; remaining comparisons were of a more qualitative nature.

Not surprisingly, given that so few subjects had mastered the discrimination within the initial training session, pre-test scores achieved one week later were significantly lower than those achieved in the second pre-test in experiment 1 ($t = 4.223$, df 6, $p < 0.005$). Of the 3 subjects who had attained 100% in the first post test, two were unable to reproduce this high level of success in this second session, both having selected the 'wug' in the single trial in which it was presented as an alternative stimulus, and both repeating this selection in the subsequent post-test (see Table 7:4). The same two subjects did achieve 100% in the pre-test administered in week 3, having apparently dropped their tendency to select the 'wug'. With the exception of the single subject who had achieved 100% scores throughout trials, however, the remaining subjects were clearly still unable to master this second discrimination. Pre-test scores in this third session were again significantly lower than those achieved in the third pre-test in study 6a ($t = 3.333$, df (6), $p < 0.01$). Even after a third training session, two subjects still showed a confusion between the previous and the present target stimuli.

DISCUSSION

Study 6b aimed to investigate whether the errorless learning procedure has any intrinsic learning value in terms of the transferability of skills learned. Results from study 6a demonstrated that a discrimination taught using the procedure was reliably reproduced up to 6 weeks following initial training but results from this second experiment were considerably less encouraging. The majority of subjects, despite responding equally well in initial training with the new target stimulus, were unable to maintain the high level of subsequent pre- and post-test performance demonstrated in study 6a. A small number of subjects moreover had severe difficulty making the transfer within training itself. In two cases sessions had to be abandoned, subjects having become 'stuck' on several trials, repeatedly selecting an incorrect figure. Unlike study 6a, it was not possible in these cases to rectify these errors through re-presentation of the previous trial; often a pass on initial presentation would be followed by a fail when the same trial set was returned to a second time. In many cases these failures did not appear to be genuine but rather to reflect an almost defiant refusal to comply with the procedure. This represents a strong contrast to children's initial responsiveness to the first training in study 6a. The carry-over effect within training scores also, in some cases, disappeared, with subjects attaining lower scores in second and even third sessions than in the first.

Why should the procedure which had proved to be so effective in the first experiment produce such comparatively poor results with this second task? Given that in both experiments training procedures had been identical, with alternative stimuli 'faded in' on the basis of size, it seems very unlikely that the second discrimination was intrinsically any more difficult to master than the first.

A particularly prevalent characteristic of performance which emerged in the first post-test of some subjects and was still in evidence in some final sessions may provide an explanation. In many cases, although correctly selecting the 'nim' in 6 out of 7 post-test trials, subjects showed a clear preference for the 'wug' in the single trial set in which it was presented. For some reason the original discrimination was given priority over the new target stimulus. Previous research with the fading technique has suggested that it does not provide sufficient comparative experience to permit transfer (Gollin and Savoy 1968) because it tends to confine subjects' attentional responses to specific attributes of the target stimulus. Similarly Bijou (1977) and Schilmoeller and Etzel (1977) pointed to the overuse of non-criterion related cues as an explanation for poor transfer (in the present study this would be the size of the target stimulus). The predominant focus on successful responses in errorless procedures is such that subjects are not required to identify the characteristics of any alternative stimulus in order to distinguish it from the target stimulus. Selection of the target stimulus is guided in initial trials through fading, and through recognition in later trials. It is possible therefore to achieve 100% success in training without necessarily having to attend to the alternative stimuli; the procedure can teach subjects how to identify which of the three figures is the correct choice, but not why the remaining two are incorrect.

On this basis, it could be argued that although the second training task enabled identification of the 'nim' as the new target stimulus, by its nature the errorless procedure precluded the possibility of learning that the 'wug' now represented an incorrect response. The identical nature of the training task and the occasional appearance of the 'wug' may have implied that the original stimulus was still relevant in some way. If children were simply responding to the non-criterion size cue in training in both tasks, it is not difficult to see how they could have become confused between the new stimulus and the original stimulus. This interpretation would suggest that rather than fully learning a discrimination skill, children were learning a task-specific strategy for making correct responses.

It will be recalled from chapter 6 that this task-specific strategic approach had been offered as a possible explanation for Ss' apparent failure to consolidate newly acquired skills. DS Ss frequently seemed to acquire the skills required to pass certain BSID items in isolation, demonstrating little evidence that they understood the association between these and other related skills. Further evidence from a study of object concept development in DS children would also seem to indicate that this tendency to adopt superficial task strategies is characteristic of the approach to learning in other areas of development (Wishart 1990). Rather than using what has been learned at one level to bridge the step to a more advanced level, each task is tackled anew. The inappropriate transfer of tactics from one task to another also seems to imply that very little true learning is actually taking place. Without actually having learned the association between the two tasks, the response is to use the strategy available at that particular time.

In the present experiment, the same strategy *was* appropriate for both tasks but the response learned through use of this strategy in the first task was not appropriate in the second task. This use of task-specific strategies lends further support to Morss' suggestion that learning in DS may be incomplete. Given that this contention was based on childrens' apparent failure to appreciate the significance of error in learning situations, it indicates a need for extreme caution in the use of techniques such as errorless learning which actually preclude erring. If DS children are naturally devising a series of superficial tactics for minimising failure, overuse of techniques which operate on similar principles may simply reinforce their belief in the efficacy of this approach.

It was suggested at the very beginning of this thesis and in the introduction to this chapter however, that DS children, because of the tendency to avoid learning situations in which they could potentially fail, may require an initial 'boost' before they can become sufficiently confident to deal with failure. Even if the use of success-only techniques does encourage the adoption of task-specific strategies, such techniques could nevertheless provide a 'baseline' of repeated success, generally not as readily achievable in conventional trial-and-error learning situations. Experience of repeated success has been demonstrated here both to improve children's performance on a task that did involve the possibility of failure, and to increase the likelihood of reliable production of correct responses in the long term. Although results from the present study imply that what is 'learned' through use of the errorless approach is of little value in terms of discrimination learning per se, in terms of what children may learn in relation to the experience of learning itself, this approach may yet be of considerable potential.

On this argument the intrinsic value of errorless learning may be not in its use as a teaching device itself, but as a device for persuading children to perform to full potential. The approach of enhancing performance on conventional learning tasks could be incorporated into an ongoing learning situation in which erorrless and trial-and-error methods are used alternately. In discrimination learning for example, having learned through errorless training which is the correct response and to reliably and consistently select that response, the child could then be taught to relinquish the particular strategy used in the choice of that response and to learn why it is in fact correct. It may be possible in this way to reduce the influence of motivational deficits on the approach to learning in DS children and to increase the stability of that performance, thereby allowing what is learned to be adequately consolidated.

By increasing reliability and stabilising performance it would seem that certain procedures of learning do produce performances which more closely match levels of performance in DS children. Although the errorless procedure was investigated here for use in teaching discrimination skills only, results imply that by increasing the experience of success/failure during skill acquisition, it may be possible to encourage a more equal relationship between competence and performance in the development of cognition in DS.

CHAPTER 8

CONCLUSIONS

Aims:

The aims of this thesis were four-fold. The major aim was to study the early development of cognitive abilities in DS from a perspective which emphasises the contribution of non-cognitive factors - in particular motivation - to the process of cognitive development in this population of children. In doing so it was hoped to bring the study of development in DS more in line with current mainstream developmental theory in which it is now largely accepted that cognitive development cannot be studied in isolation but should be seen as a complex and dynamic process, a process which is influenced as much by what the child is motivated to learn as by his/her level of cognitive potential. Until recently the study of development in DS and in other mentally handicapped children has been heavily dominated by an approach which focuses primarily on the functional deficits and delays which distinguish the child with a mental handicap from the normally developing child. This thesis has aimed to provide a broader perspective on the learning difficulties experienced by the DS child by considering the way in which such functional deficits might have influenced his/her approach to learning.

A second and related aim was to demonstrate that poor performance cannot always be taken to indicate similarly poor levels of cognitive competence in the DS child and that motivational deficits may impede the demonstration of optimal levels of competence. From this approach it was aimed to determine the extent to which any competence/performance differential might affect the performance of infants and young children with DS in assessment situations and in the light of these findings, to re-evaluate those theories of development in DS which have been based solely on outcome measures derived from psychometric assessments.

A third aim was to explore the possibility that motivational problems may inhibit the developmental process itself, thereby adding to the already existing cognitive deficit. The idea that poor motivation can adversely affect the performance of people with a mentally handicapping condition is not a new one. Many previous research studies have resulted in descriptions of this population as 'learned helpless' (Floor and Rosen 1975; Weisz 1982) or 'failure avoiders' (Cromwell 1967). Until recently, however, such descriptions have been applied from a perspective which sees the motivational deficit to be another 'symptom', inherent in the condition of mental handicap. This thesis has aimed to show, by contrast, that the motivational deficit in DS has its own developmental history and that its effects on the process of cognitive development may originate at a very early stage in development.

The final aim of this thesis was to explore how differential experience of success/failure in a learning situation might influence both the acquisition and consolidation stages of the learning process in DS children. The possibility that the true value of errorless learning teaching strategies may lie in their potential to increase the DS child's motivation to learn from more conventional teaching methods was investigated. These methods are widely used as teaching methods with the mentally handicapped, in place of traditional trial and error teaching techniques. Here it was aimed to demonstrate that a purely errorless learning approach may be inappropriate in the longer term while also showing that there may be considerable value in a teaching strategy which incorporates both errorless and trial-and-error methods.

The relationship between competence and performance in DS children

In the course of this thesis a great deal of evidence was gathered from both learning and assessment studies of DS children which lent support to the hypothesis that the relationship between competence and performance in DS is not a straightforward one. Repeatedly it was found that levels of performance fell short of optimal levels of ability. A comparison of two methods of teaching discrimination to DS and to non-handicapped children (study 1) indicated that DS performance was affected to a comparatively greater extent by prior experience of success and failure; DS subjects' performance on a trial-and-error task was significantly enhanced when this task had been preceded by an errorless learning task, but performance on both tasks was adversely affected if the two tasks were presented in the opposite order. Non-handicapped children showed no such order effects; analysis of their scores indicated that any improvements made were as much attributable to the effects of practice as to any possible enhancing effects of prior experience of learning in the absence of error.

Findings from detailed cross-sectional studies of the performance of children with DS on the BSID also revealed a strong tendency to underperform (studies 2 and 3). Subjects frequently failed items which their parents reported they should have passed. As age increased so did the likelihood that children would 'fail to engage' in test items - either by attempting to divert E's attention away from the task in hand or by refusing to comply with E's requests to demonstrate specific skills with the test objects. In study 3 it was revealed that children performed to full competence in *neither* of two closely-spaced sessions with optimal scores (based on all passes attained in either of the two sessions) significantly higher than those recorded for either session alone. Such findings strongly suggest that motivational factors adversely affect the approach of DS children to cognitive tasks.

It is well established that motivation plays an important role in normal cognitive development. Studies conducted by Dweck (1975, 1978, 1982)) have shown that learning in non-handicapped children may be adversely affected by poor levels of motivation. It has also been proposed that the assessment performance of normally developing children necessarily includes a measure of motivational levels (Hrncir 1985; Scarr 1981). As stated above, however, although it has been acknowledged that motivational factors may impede the performance of mentally handicapped children it is often assumed that the mentally handicapped child is inherently poorly motivated to learn. Theoretical approaches to the learning difficulties encountered by mentally handicapped children have therefore tended to neglect the importance of motivational factors, focussing instead on more cognitively-based explanations of their difficulties in learning situations.

The competence/performance differential: its implications for learning

The popularity of errorless learning teaching strategies for use with mentally handicapped children demonstrates the extent to which theoretical approaches to learning in mental handicap have diverged from current mainstream developmental theory. Generally the use of errorless techniques stems from the belief that mentally handicapped people cannot learn by trial and error - that the experience of erring in a learning situation will always be counterproductive. This idea has persisted in spite of the fact that it is incompatible with the idea, largely accepted in mainstream learning theory, that error has as much to contribute to the process of learning as success. Implicit in the proposal that handicapped people cannot learn from error therefore is the suggestion that the way in which they learn must be inferior in some way.

Indeed a large majority of the findings presented in this thesis could be interpreted to lend support to this suggestion. In study 4 detailed examination of longitudinal BSID profiles revealed that many children failed on items which they had passed at earlier age levels. In the majority of such cases subjects did not reliably demonstrate any higher level skills indicating that they had 'surpassed' these items, nor was it possible to fully explain their disappearance in terms of the lack of suitability of the test items. Many subjects genuinely appeared to have lost skills which they had demonstrated to be within their repertoires several months earlier - indicating that there may be associated problems with consolidation in DS. This pattern contrasts markedly with the incremental process believed to characterise normal development.

In normal development it is believed that the experience of error may in fact frequently provide the child with the incentive to build on what has already been learned. Having acquired a particular skill which proves to be effective in certain situations the child then goes on to explore this skill and to apply it to different tasks. The discovery that this leads to error leaves the child with the challenge of finding a more effective solution - which often marks the step onto a higher level of development.

Results from studies 2 and 3 indicated that the DS child, by comparison, does not seem to explore and experiment to the same degree. Observations made in study 2 suggested that rather than being motivated to explore their environment, young DS babies may actually lose the motivation to do so. It was noted, for example, that several 6 month old DS children had apparently lost their interest in locating the source of a sound. Parents provided reports supporting this interpretation of this developmental 'loss'. The responses of several 3 month old DS babies to sound-locating tasks indicated that although apparently having the motivation to turn towards the source of a sound, their ability to do so was restricted by the tonic neck reflex. This finding suggested that there may have been a causal link between these two patterns of response: that the mismatch between the desire or intention to explore and the inability to do so at the earlier age level had resulted in the loss of interest seen in older DS children. Cunningham reported similar findings from a study of visually directed reaching in DS babies of 16 to 18 weeks of age.

Results from study 2 indicated that a similar history of mismatching intentions and abilities might explain the frequency with which older DS children were avoiding BSID tasks. It was revealed that several items which children were avoiding were the same items which had been failed by large numbers of children at earlier age levels. Detailed investigation of the longitudinal profiles of three individual children (study 5) revealed that task avoidance did indeed seem to be related to failures encountered on similar tasks during testing sessions 3 and 6 months previously. A link between unreliable performance on certain items and the later loss of the same items emerged - a link which strongly suggested that there may have been a causal relationship between poor reliability during the acquisition stages of learning and subsequent failure to consolidate. Additional observations, moreover, lent support to the hypothesis that this pattern of task avoidance and subsequent consolidation difficulties might have compromised children's understanding of the successes they did achieve. Such findings lend support to Morss' (1979) suggestion that learning in DS may be incomplete.

All of the above lends support to the suggestion that learning in DS may be adversely affected by the early experience of error encountered during the acquisition of certain skills. Such an explanation does not in itself indicate, however, that the DS child is incapable of learning from errors. On the contrary the findings of study 5 suggested that the frequency with which DS children encounter errors which cannot contribute to the overall developmental process may prevent the child from experiencing - and identifying - the sorts of errors which present a challenge to the non-handicapped child. DS children often rejected tasks after making errors which were attributable more to the motor deficit than to any inability to meet the cognitive demands of the task. This tendency to avoid error may generalise to errors which are necessary for a full understanding of success thereby leaving the DS child with only a vestigial understanding of the concepts required for development to progress normally.

Encouraging learning in the absence of error may not therefore be the most effective means of overcoming the learning difficulties experienced by the DS child. Indeed such an approach may actually accentuate the tendency to avoid error. A more valuable approach may instead involve altering the DS child's success/failure experience ratio in a learning situation in an attempt to reduce the probability that that learning situation will be avoided. The results of study 1 indicated that errorless learning techniques may have a useful contribution to make in this area. It was demonstrated that prior 'priming' with an errorless strategy could enhance DS children's performance on a trial-and-error task. Here the main purpose of the errorless task was not to teach the child a discrimination - rather it was being used to teach the child that he/she can learn - that learning need not always be such an unrewarding process.

An equally encouraging set of results emerged from study 6a in which it was demonstrated that discrimination skills taught using the errorless teaching procedure were retained and reliably reproduced by DS children for up to 6 weeks. This result contrasts sharply with data obtained from the BSID studies in which DS children's performance on test items was seen to change both in the short term and in the long term. Not only does this result imply that children are less likely to avoid tasks taught using the errorless method, but also that skills taught using this method may become more stable acquisitions to the developmental repertoire.

Unfortunately a subsequent experiment (study 6b) revealed that such 'skills' may have been no more than task-specific responses. In this study DS Ss were found to be unable to transfer the newly acquired discrimination 'skill' to a second discrimination task. Detailed examination of response patterns moreover indicated that the poor levels of success seen may have been attributable to Ss' inability to distinguish between the correct responses to the two tasks. The errorless strategy had provided no means by which to 'unlearn' the response taught in the first task.

Although this finding does have very negative implications in relation to the effectiveness of success-only methods for teaching specific skills, it does demonstrate that the experience of error in learning is as important for the child with DS as it is for the non-handicapped child. This set of results indicates the need for a more global approach to the learning difficulties experienced by the DS child - one which considers both cognitive and non-cognitive factors. Rather than simply accepting that the experience of error in a learning situation will always be counterproductive, such an approach demands an explanation of the developmental role of errors in DS learning.

Implications for development: the developmental difference theory

The pattern of early development in DS which has emerged from the studies presented in this thesis contrasts sharply with the incremental developmental pattern seen in the non-handicapped child. It has been proposed that development in DS is characterised by frequent failure, poor rehearsal of new acquisitions as a result of the tendency to avoid experience of failure and, in turn, poorly developed levels of understanding and difficulties with the consolidation of newly learned skills.

There was little indication that this pattern of development was restricted in DS children to any particular skill areas. Results from chapters 3 and 4 did reveal that the tasks on which children were producing unreliable performances and those in which they were failing to engage were similar in type. Study 5 also indicated that failure to engage was often associated with previous experience of failure on similar tasks. It could therefore be stated that there are particular weaknesses associated with the condition of DS in relation to these specific skill areas. Results from chapter 6, however, revealed very little in the way of evidence to suggest that these specific weaknesses would predictably lead to the loss of these same skills in all subjects. The relationship between unreliability and instability was identified across a wide range of test items - often in single cases only. Therefore, although at some level it could be stated that development in DS is associated with poor reliability of performance which may lead to developmental instability, it is not possible to specify how this sequence may relate to the overall pattern of development. It is not, in other words, possible to determine with any confidence how this 'different' pattern of development will be manifested in terms of the repertoire of emerging abilities in the individual DS child.

It may therefore be more useful to attempt to explain the pattern of development in DS which has emerged from these studies at a more global, theoretical level - one which focuses less on specific patterns of cognitive or motor skills and abilities and more on the way in which the development of these abilities may be affected by factors other than levels of cognitive or motor functioning. At this level, rather than identifying differences, it is possible to see how factors influencing development in DS are very similar to those believed to characterise normal development. Although this thesis has highlighted the motivational deficit in the DS child it has nevertheless drawn attention to the fact that motivational factors cannot be left out of any theory of cognitive development in DS. As in normal development cognition and motivation are closely interrelated in DS development.

With the emphasis on such similarities in the overall structure of development, rather than on the differences, delays or aberrations observed in very specific areas, it is possible to conceive of a single theory which might encompass both patterns of development. This more integrated theoretical approach could not only provide a means by which mainstream developmental ideas could be applied to mental handicap, but it is also likely to benefit the DS child in many positive ways.

This thesis is not the first attempt to look at development in DS in this more global way. Proponents of what is termed the 'organisational' approach have attempted to demonstrate that the interrelationships between cognition and other developmental domains, for example, affect, are very similar in DS and normal development (see e.g Beeghly and

Cicchetti 1987). Unfortunately the value of much of this research has been diminished by its reliance on psychometric data for matching DS and non-handicapped children. The work reported in this thesis strongly suggests that psychometric instruments cannot provide accurate measures of the cognitive abilities of children with DS. At one level it is possible to see in the organisational approach a rejection of the medical model which concentrates solely on the 'symptoms' of mental handicap and the ways in which these interfere with the normal process of development. At another level, however, it could be argued that the medical model of mental handicap is being perpetuated by the use of the very tools which emphasise such symptoms: psychometric-based assessment batteries.

A similar dichotomy can be identified in the work of Rauh et al (1990 in press) aimed at measuring the rate of development in DS. On one hand, this group of researchers have conceded from their results that there is no prescribed rate at which development in DS progresses and have concluded that it may be more productive to consider the shape of the developmental curves of individual children. On the other hand, however, by using psychometric tools to measure DS development, these studies could only result in a negative picture of development in DS. If developmental rate is measured only against the norm, the outcome will inevitably highlight the fact that DS children take longer to acquire specific cognitive skills. It cannot accurately monitor any positive progress which the DS child may have made over time.

Implications for psychometric assessment

As Cicchetti (1982) pointed out, the reliance on psychometric measurements may in large part be attributable to the fact that no alternative means of assessing development in mental handicap are as yet available. The only available frame of reference is that which has come from years of studying normal development.

The original objective in compiling assessment batteries such as the BSID was to devise a means of determining the extent to which individual differences in intelligence remained constant from infancy to school age. Despite all efforts to maximise numerical precision, however, it failed to generate any evidence to support the predictive validity of infant 'IQ'. The BSID was consequently redesigned to serve more diagnostic purposes. Bayley (1969) stated that the primary value of the revised version of the BSID is that it provides "the basis for establishing the child's current developmental status and thus the extent of any deviancy from normal expectancy" and "a basis for instituting early corrective measures when the child shows evidence of retarded mental.....development". Having failed in its original objective the BSID therefore became a 'hand-me-down' for use with populations of children who were not represented in the standardisation sample. No consideration was given to the possibility that it might not be appropriate for this purpose and no attempts were made to adapt the scale for use with handicapped children.

There are, however, a number of reasons for questioning the concurrent validity of the BSID when used with DS children. The BSID has a system of scoring which can only accurately measure the 'developmental status' of children who fall within the upper ranges of ability in mental handicap: the MDI stops at 50. In addition, the test itself contains a large number of items which have a heavy motor component and which therefore cannot hope to assess accurately the cognitive abilities of any child who has a deficit in motor ability. Items were originally designed for use with children whose chronological ages fell between 0 and 30 months and yet it was simply assumed that these same items would be suitable for use with handicapped children of considerably older ages. More importantly perhaps, the step sizes between differing levels of difficulty in the BSID were calculated on the basis of the progress made by non-handicapped children over fixed time intervals. It is consequently unable to monitor any smaller-scale developmental attainments which handicapped children may make over the same time periods with the BSID. Its validity for use with DS children is further compromised by the fact that this group of children tend to avoid many cognitive tasks.

This test also cannot claim to be a reliable means of measuring the cognitive development in children with DS. The performance of DS children on the BSID was found to be considerably more variable than would be expected on the basis of normative reliability studies. Study 3, moreover, revealed that the level of reliability quoted for the BSID may actually be artefactually low - thereby indicating an even greater difference between DS and non-handicapped children in the extent to which they can be expected to produce reliable BSID performances.

Studies of the predictive validity of infant tests such as the BSID have resulted in the conclusion that there is very little correlation between individual non-handicapped children's infant and childhood IQs but that this correlation is higher for children with mental handicap (Illingworth and Birch 1959; Knobloch 1959; Knobloch and Pasamanick 1960, 1963, 1967; Koch et al 1963). This, it has been suggested, indicates that such tests can indeed have some predictive function within such populations. Study 4 casts considerable doubt on this claim. Given the level of item instability between adjacent pairs of ages found in that study, it seems highly unlikely that this claim, made on the basis of overall scores, would be substantiated if the same data which led to this claim had been analysed on an item-item basis.

From all of the above it is clear that the BSID - and other psychometric tests of infant development - can provide only very limited information on the development of children with DS. What it can tell us in relation to this - and many other groups of handicapped children - may well be inaccurate and is predominantly negative in nature. Clearly there is a pressing need for a better and more positive means of assessing the abilities of the child with mental handicap. This need seems particularly pressing given the latest set of guidelines produced by the government in relation to the assessment of mentally handicapped individuals. The most recent note of guidance on the implementation of the proposals laid down in the 1987 White Paper on community care is devoted entirely to the subject of assessment. The very existence of this publication implies a strong commitment to formal methods of assessing the needs of the mentally handicapped population. Such an emphasis on assessment may have extremely unfavourable consequences for the mentally handicapped child. If there are no viable alternatives to the psychometric measures currently in use it is likely that such measures will be officially adopted for the purpose of assessing the early intervention needs of this age group of the mentally handicapped population. Clearly these measures are unlikely to provide an accurate picture of the needs of DS children or of their potential to benefit from appropriately tailored intervention.

As Garwood (1982) has pointed out, the curriculae for many intervention programmes are based on the sequences of items contained in psychometric tests. This implies that many of the inadequacies discussed throughout this thesis in relation to the use of such tests in assessing mentally handicapped children also apply to the methods of intervention used with such children. The Portage intervention programme, widely used by home teachers in this country, uses a checklist of items which are matched to curriculum cards. Teaching is focussed on items on the checklist which are absent from the handicapped child's profile - the aim being to provide experiences designed to facilitate the emergence of the skills required to pass such items. Not only does this programme implicitly assume that children develop behaviours in the invariant order in which they appear in the checklist, but also that the process of development consists only of the acquisition of the very specific behaviours contained on that checklist. The overall objective of the Portage programme is to determine what is 'missing' from the behavioral repertoire of the handicapped child and then to attempt to 'patch up' these gaps by focussing on the behaviours themselves. It is simply assumed that their absence is attributable solely to the functional deficits associated with the handicapping condition - little consideration is given to the influence - whether positive or negative - of motivational factors on development. This 'patching' may be successful at some level but it fails to address the basic problem that mentally handicapped children may not be as motivated to learn as their non-handicapped peers.

In view of the above, it is perhaps not a surprise that despite the effort and funds presently being diverted into early intervention programmes of this sort for DS children, the ability levels found in the studies in this thesis differ very little from those seen in earlier studies of this population. Clearly, if neither intervention programmes nor instruments for assessment are sensitive to the needs of children with a mental handicap, it is likely that both the programmes and studies evaluating the efficacy of these programmes will continue to produce disappointing outcomes. Unless it is stressed, however, that this lack of improvement in levels of ability may simply reflect the inadequacy of currently available programmes for early intervention and of methods of assessment of progress, such findings could well be interpreted as lending support to the argument that there is little potential for improvement in the DS child.

Future directions

There is a need for the application of a more integrated approach both to clarify the assessment and intervention needs of the DS child. By highlighting the influence of motivational factors in the development of cognition in DS it is hoped that this thesis has demonstrated the merits of an approach to the learning difficulties in this population which considers the whole child and not just his/her level cognitive functioning in isolation. Research should continue to look beyond the deficits and delays which are identified through comparison with normally developing children and focus instead on attempting to understand the complex interactions between cognitive, motivational, social and emotional factors which might explain why it is that the DS child experiences difficulties in learning. It is no longer sufficient simply to attribute all of the difficulties which emerge to the handicapping condition.

A more accurate picture of development in DS might best be achieved by studying DS children in their own right. It may be of more value to focus simply on what DS children themselves can do rather than to measure their achievements from within the parameters of normal development. Such an approach may not only provide a more sensitive insight into the way in which the DS child does learn, but may also lead to a very much more positive outlook on the DS child than that obtained through constant reference to his/her more able peers. Where it is necessary to provide an assessment measure, rather than pointing out that a child has an IQ of less than 50 it may be of more value to stress what he or she can do - that s/he can pass the Piagetian stage 5 object concept tasks and can make sentences of three words, for example. In a related vein, in research studies where it is required to match DS subjects with non-handicapped subjects, a similar approach to that adopted in the errorless learning study presented in Chapter 2 might be taken. In that study, it will be recalled, subjects were matched not on the basis of equivalent mental ages but on the basis of the specific discrimination skills required for participation in the experiment.

A more positive approach to the motivational deficits evidenced by DS children would also seem to be merited. Evidence was presented in study 2, for example, suggesting that the auditory discriminatory abilities of the DS child may be affected at a very young age by poor levels of motivation. Success-only or fading techniques may have a useful role to play at this early stage in development by providing the child with a sufficiently high experience of success to maintain the will to explore the environment. It would also seem well worth exploring the efficacy of the errorless learning approach for teaching skills other than discrimination. However, it would be equally valuable to identify situations in which DS children are motivated to respond to their own errors i.e. to focus as much on their strengths as on the weaknesses associated with the condition.

It is also important that while remaining aware of the differences in development between handicapped and non-handicapped children, we continue to stress the similarities between the two populations. The successful integration of any minority group into our society has worked on this principle. The mentally handicapped are not ill - even government policy has acknowledged this issue through moves towards the closure of hospitals for this population. For the Care in the Community initiative to succeed it is vital that we continue to point out the inadequacies of the medical model of handicap and to emphasise the rights of the handicapped as individuals - in terms of research attention and funding as well as in standards of care and educational provision. Research in particular must attempt to understand why it is that competence is not more efficiently reflected in performance in DS children and identify methods of ameliorating this differential.

CHERKES-JULKOWSKI, M., GERTNER, R., and NORLANDER, K. (1986) Differences in cognitive processes among handicapped and average children: a group learning approach. Journal of Learning Disabilities, 19, 438-455.

CICCHETTI, D. and BEEGHLY, M. (Eds.) Children with Down syndrome: A developmental perspective. Cambridge University Press 1990.

CICCHETTI, D. and GANIBAN, J. Developmental processes in infant Down Syndrome. In R. M. Hodapp, J. A. Burack and E. Zigler (Eds.) Issues in the development approach to retardation. Cambridge University Press 1990.

CICCHETTI, D. and POGGE-HESSE, P. Possible contributions of the study of Organically retarded persons to developmental theory. In E. ZIGLER and D. BALLA (eds), Mental Retardation: The Developmental-Difference Controversy. Hillsdale, New Jersey: Erlbaum, 1982.

CICCHETTI, D. and SCHNEIDER-ROSEN, K. 1983: Theoretical and empirical considerations in the investigation of the relationship between affect and cognition in atypical populations of infant. In C. IZARD, J. KAGAN and R. ZAJONC (Eds.) Emotions, Cognition and Behaviour. New York: Cambridge University Press.

CICCHETTI, D. and SERAFICA, F. (1981) The interplay among behavioural systems: Illustrations from the study of attachment, affiliation and wariness in young Down's Syndrome children. Developmental Psychology 17, 36-49.

CICCHETTI, D. and SROUFE, L. A. An organisational view of affect: Illustration from the study of Down Syndrome infants. In M. LEWIS and L. ROSENBLUM (Eds), The Development of Affect. New York: PLenum 1978.

CICCHETTI, D. and SROUFE, L. A. The relationship between affective and cognitive development in Down's Syndrome infants. Child Development, 1976, 47, 920-929.

CLARKE, A. D. B., and CLARKE A. M. (1975). Recent Advances in the Study of Subnormality. London: Mind.

CLARKE, A. D. B., and CLARKE, A. M. (1973a) Mental Retardation and Behavioural Research IRMR Study Group 4 Edinburgh: Churchill Livingstone.

CLARKE, A. M. and CLARKE, A. D. B. (1974) Mental Deficiency: The Changing Outlook. (3rd Ed) London Methuen.

CLEMENTS, P. R., BATES, M. V. and HAFER, M. (1976). Variability within Down's syndrome (trisomy-21); empirically observed sex differences in IQs. Mental Retardation, 14(1), 30-1.

COHEN, L. B. (1981) Examination of habituation as a measure of aberrant infant development. In S.L. Friedman and M. Sigman (Eds.), Preterm birth and psychological development (pp. 241-253). New York: Academic Press.

CONGER, J. A. (1930). An evaluation of the Linfert scale for measuring the development of infants. Unpublished Master's thesis, University of Minnesota.

CORNWELL, A. C. and BIRCH, H. G. (1969). Psychological and social development in home-reared children with Down's Syndrome (mongolism). American Journal of Mental Deficiency 75 (3), 341-50.

COWIE, V. A. (1970). A study of the early development of mongols. Oxford: Pergamon Press.

COX, A. RUTTER, M., YULE, B. and QUINTON, D. (1977) Bias resulting from missing information: Some epidemiological findings. British Journal of Preventative and Social Medicine, 31, 131-136.

CRASKE, M. L. (1988) Learned helplessness, self worth motivation and attribution retraining for primary school children. British Journal of Educational Psychology, 58, 152-164.

CROMWELL, R. L. A Social Learning Approach to Mental Retardation. In ELLIS, E. (Ed) Handbook of Mental Deficiency. New York: McGraw-Hill 1963.

CROMWELL, R. L. Success-Failure Reactions in Mentally Retarded Children. In J. ZUBLIN and G. A. JERVIS (Eds) Psychopathology of Mental Development New York: Grune and Stratton (1967).

CULLEN, C. (1976) Errorless Learning with the Retarded . Nursing Times vol 7225th March p 45-47.

CUNNINGHAM, C. and MCARTHUR, K. (1981). Hearing loss and treatment in young Down's Syndrome children. Child: Health Care and Development 7, 357.

CUNNINGHAM, C. C. (1974) The Relevance of 'Normal; Educational Theory and Practice for the Mentally Retarded. In J. TIZARD (Ed) Mental Retardation: Concepts of Education and Research (IRMR Symposia 12-14) London: Butterworth.

CUNNINGHAM, C. C. (1979). Aspects of early development in Down's Syndrome Infants. PhD Thesis, University of Manchester.

DAVIES, B. (1985) Hearing Problems in B. Stratford and D. Lane. (Eds.), Current Approaches to Down's Syndrome.

DAVIS, J., ROBERTSON, F. M., CULLEN, C., (1975) Learning Colours Apex Journal of the Institute of Mental Subnormality 1975 2 (4) 14-15, 3 (1) 14-15.

de CORIAT, L. F., THESLENCO, L. and WAKSMAN, J. (1967) Proceedings of the First Congress of the International Association for the Scientific Study of Mental Deficiency. Eds B. W. Richards, Michael Jackson, Reigate, p 377-385.

DELOACHE, J. S., RISSMAN, M. W. and COHEN, L. B. (1978) An investigation of the attention-getting process in infants. Infant \Behaviour and Development, 1, 11-25.

DEPARTMENT OF HEALTH AND SOCIAL SECURITY (1983), Health and Service Department, Care in the Community and Joint Finance, HC (83) 6, DHSS, London.

DICKS-MIREAUX, M. J. (1966) Development of intelligence of children with Down's Syndrome: a preliminary report. Journal of Mental deficiency Research, 10, 89.

DICKS-MIREAUX, M. J. (1972). Mental development of infants with Down's Syndrome American Journal of Mental Deficiency 77 (1), 26-32.

DIENER, C. and DWECK, C. S. (1980) An analysis of learned helplessness: I The processing of success. Journal of Personal and Social Psychology 39, 940-952.

DONNAI, D. and ANDREWS, T. Screening for Down's syndrome. British Medical Journal 1988.

DOWN, J. L. (1866) Observations on an ethnic classification of idiots. Clinical Lectures and Reports, London Hospital 3 259.

DUFFY, L. (1985) Comparison of Two Procedures for Teaching Discrimination in Down's Syndrome and Normal Children. Unpublished MA Honours Thesis, Department of Psychology, University of Edinburgh.

DURHAM, M., and BLACK, K. N. (1978) The test performance of 16-21 month olds in home and laboratory settings. Infant Behaviour and Development, 1, 216-223.

DWECK, C. (1975) The role of expectations and attributions in the alleviation of learned helplessness. Journal of Personality and Social Psychology, 31, 674-685.

DWECK, C. S (1986) Motivational processes affecting learning. American psychologist, 41 (10), 1040-1048.

DWECK, C. S. and BUSH, E. S. (1976). Sex differences in learned helplessness: I. Differential debilitation with peer and adult evaluators. Developmental Psychology, 12, 147-156.

DWECK, C. S. and REPUCCI, D. (1973). Learned helplessness and reinforcement responsibility in children. Journal of Personal and Social psychology, 25, 109-116

DYER, S., GUNN, P., RAUH, H. and BERRY, P. Motor development in Down Syndrome children: an analysis of the Motor Scale of the Bayley Scales of Infant Development. Paper presented at the IASSMD Congress, Dublin, Ireland, August.

EIPPER, D. S., and AZEN, S. P. (1978) A comparison of two developmental instruments in evaluating children with Down Syndrome. Physical Therapy, 58, 1066-1069.

ELLIS, N. R. (1969) A Behavioral Research Strategy in Mental Retardation: Defense and Critique American Journal of Mental Deficiency 73 557-66.

ETZEL, B. C., LeBLANC, J. M., SCHILMOELLER, K. J., and STELLA, M. E. (1981). Stimulus control procedures in the education of young children. In BIJOU, S. W., and RUIZ, R. (Eds.), Behaviour Modification: Contributions to Education. Hillsdale, NJ: Erlbaum.

EVANS, J. A., HUNTER, A. G. W. and HAMERTON, J. L. (1978). Down syndrome and recent demographic trends in Manitoba. Journal of Medical Genetics 15, 43-7.

FABIA, J. and DROLETTE, M. (1970). Life tables up to age 10 for mongols with and without congenital heart defect. Journal of Mental Deficiency Research, 14, 235.

FANNING, G. S. (1971). Vision in children with Down's Syndrome. Australian Journal of Optomology. 54, 74-82.

FANTZ, R. L., FAGAN, F. F., MIRANDA, S. B. (1975). Early visual selectivity > In L. B. Cohen and P. Salpatek (Eds.), Infant perception: From sensation to cognition, Vol 1, Basic visual processes (pp. 249-346), New York: Academic Press.

FISHLER, K. (1975). Mental development in mosaic Down's syndrome as compared with trisomy-21. In R. KOCH and F. F. de la CRUZ, Down's syndrome (mongolism): Research, prevention and management. New York: Bruner/Mazel.

FISHLER, K., SHARE, J. and KOSH, R. (1964) Adaptation of Gesell developmental scales for evaluation of development in children with Down's Syndrome. American Journal of Mental Deficiency, 68, 624-6.

FOGELMAN, K. (1985). Exploiting longitudinal data: Examples from the National Child Development Study. In NICOL, A. C. (Ed.) Longitudinal Studies in Child Psychology and Psychiatry. John Wiley and Sons Ltd.

FRAIBERG, S. 1977: Insights from the blind. London: Souvenir Press.

FRANCES, P. L., SELF, P. A., and HOROWITZ, F. D. (1987). The behavioural assessment of the neonate: an overview. In OSOFSKY, J. D. (Ed) Handbook of Infant Development (2nd Ed.) New York: Wiley.

FRAZER, J. and MITCHELL, A. 'Kalmuc idiocy: report on a case with autopsy and notes on 62 cases'. Journal of Mental Science, 22, 169-72.

FUCHS, D., FUCHS, L. S., DAILEY, A. M. and POWER, M. H. (1985a) The effects of examiners' personal familiarity and professional experience on handicapped children's test performance. Journal of Educational Research, 78, 141-146.

FUCHS, D., FUCHS, L. S., POWER, M. H. and DAILEY, A. M. (1985) Bias in the assessment of handicapped children. American Educational Research Journal, 22, 185-198.

GALLAGHER, R. P. and LOWRY, R. B. (1975). Longevity in Down's Syndrome in British Columbia Journal of Mental l Deficiency Research 19, 157.

GARDINER, P. (1967). Visual defects in Down's and other mentally handicapped children. British Journal of Opthalmology, 57, 469-74.

GARDNER, W. I. Effects of Interpolated Success and Failure on Motor Task Performance in Mental Defectives. Paper presented at the meeting of the Southeastern Psychological Association, Nashville Tennessee March 1957.

GARWOOD, G. S. (1982) (Mis) use of developmental scales in program evaluation. Topics in Early Childhood Special Education, 1, (4), 61-69.

GATH, A. (1985). Down's Syndrome in the first nine years. In NICOL, A. C. (Ed) Longitudinal Studies in Child Psychology and Psychiatry. John Wiley and Sons Ltd.

GESELL, A., and THOMPSON, H. The Psychology of Early Growth. New York: MacMillan, 1937.

GIBSON, D. (1978) Down's Syndrome: The Psychology of Mongolism. Cambridge: Cambridge University Press.

GILLILAND, A. R. The measurement of the mentality of infants. Child Development, 1948, 19, 155-158.

GIRARDEAU, F. L. (1959) The Formation of Discrimination Learning Sets in Mongoloid and Normal Children. Journal of Comparative and Physiological Psychology 52 566-70.

GLENN, S. M., CUNNINGHAM, C. C. and JOYCE, P. F. (1981). A study of auditory preferences in nonhandicapped infants and infants with Downs syndrome. Child Development, 52, 1303-1307.

GOLLIN, E. S. and SAVOY, P. (1968) Feeding Procedures and Conditional Discrimination in Children. Journal of the Experimental Analysis of Behaviour 11 443-451

GORDON, A. M. (1944) Some Aspects of Sensory Discrimination in Mongolism. American Journal of Mental Deficiency 49 55.

GREENWOOD, R. D. and NADAS, A. S. (1976) The clinical course of cardiac disease in Down's Syndrome. Pediatrics 58, 893.

GREGORY, O (1973) The Use of Piagetian Assessment Techniques in the Assessment of Profoundly Retarded Children. Paper presented to Third International Congress of I. A. S. S. M. D., The Hague.

GUIDUBALDI, J. and CHRISTIE, A. (1982) The Battelle Developmental Inventory Test Development, norming and validity studies. Paper presented at the National Handicapped Children's Early Education Program/Division of Early Childhood Conference, Washington, DC.

GUIDUBALDI, J. and PERRY, J. (1984). Concurrent and predictive validity of the Battelle Developmental Inventory at the first grade level. Educational and Psychological Measurement, 44, 977-985.

GUIDUBALDI, J., NEWMAN, I., CLEMINSHAW, H., PERRY, J. TELZROW, C., SERAZIN, R. MARANDA, M. VETTEL, M. and HARR, G. (1981) The Battelle Developmental; Inventory: Construction, norming and validation. Paper presented at the National Association of School Psychologists, Houston, Texas.

GUNN, P., BERRY, P. and ANDREWS, R. (1981). The temperament of Down's syndrome infants: A research note. Journal of Child Psychology and Psychiatry, 22, 189-194.

HAMERTON, J. L., GIANELLI, F. and POLANI, P. E. (1965). Cytogenetics of Down's syndrome (mongolism). I. Data on a consecutive series of patients referred for genetics counselling and diagnosis. Cytogenetics, 4, 171.

HANSON, J. M. (1981) Down's Syndrome children: Characteristics and intervention research. In M. LEWIS and L. H. ROSENBLUM (Eds) The Uncommon Child, p 83-114, New York: Plenum.

HANSON, M. J., and SCHWARTZ, R. H. (1978). Results of a longitudinal intervention program for Down Syndrome infants and their families. Education and Training of the Mentally Retarded, 13, 403-407.

HARTER, S. and ZIGLER, E . Effects of Stimulus Presentation and Penalty Conditions on the Discrimination Learning of Normal and Retarded Children. Developmental Psychology 1972 6 85-91.

HERRING, A. (1937) An experimental study of the reliability of the Buehler Baby Tests. Journal of Experimental Education, 6, 147-159.

HILL, P., and McUNE-NICOLICH, L. (1981). Pretend play and patterns of cognition in Down's syndrome infants. Child Development, 23, 43-60.

HINDLEY, C. P. (1960) The Griffith Scale of Infant Development: scores and predictions for 3-18 months. Journal of Child Psychology and Psychiatry, 1, 99-102.

HODAPP, R. M. and ZIGLER, E. Applying the developmental perspective to individuals with Down Syndrome. In D. Cicchetti and M. Beeghly (Eds.) Children with Down syndrome: A developmental perspective. Cambridge University Press 1990.

HODAPP, R. M., BURACK, J. A. and ZIGLER, E. (Eds.) Issues in the developmental approach to retardation. Cambridge University Press 1990.

HORNER, T. M. (1980). Test-retest and home-clinic characteristics of the Bayley scales of Infant Development in 9- and 15- month old infants. Child Development, 51, 7541-758.

HORNER, T. M. (1988) Single versus repeated assessments of infant abilities using the Bayley Scales of Infant Development. Infant Mental Health Journal, 9 (3).

HRNCIR, E. J., SPELLER, G. M. and WEST, M. (1985) What are we testing? Developmental Psychology, 21 (2) 226-232.

ILLINGWORTH, R. S. (1959). The Diagnosis of Mental Retardationin Infancy: A Follow up Study. Archives of Diseases in Childhood, 34, 269-273.

ILLINGWORTH, R. S. (1980). The Development of the Infant and Young Child: Abnormal and Normal. (7th ed.). Edinburgh: Churchill: Livingstone.

KATLICK, R., CLARK, E. B., NEILL, C. and HALLER, J. A. (1977). Surgical management of congenital heart disease in Down's Syndrome. Journal of Thoracic and Cardiovascular Surgery. 74, 206.

KEELE, D., RICHARDS, C., BROWN, J. and MARSHALL, J. (1989). Catecholamine metabolism in Down's syndrome. American Journal of Mental Deficiency, 74, 125-129.

KEISER, H., MONTAGUE, J., WOLD, D., MAUNE, D. and PATTISON, D. (1981) Hearing loss of Down's Syndrome adults. American Journal of Mental Deficiency 85, 467-72.

KEITH, J. D., ROWE, R. D. and VLAD, P. (1978) Heart disease in infancy and childhood, 3rd edition. London: Macmillam.

KNOBLOCH, H., (1959). Letter to the Editor. Paediatrics, 23, 175-178.

KNOBLOCH, H., and PASAMANICK, B., (1960). An Evaluation of the Consistency and Predictive Value of the 40 week Gesell Developmental Schedule. In C. SHAGASS & B. PASAMANICK (Eds.), Child Development and Child Psychiatry. Washington, D.C.:American Psychiatric Association.

KOCH, R., SHARE, J. WEBB, A. and GRALIKER, B. V. (1963). The predictability of Gesell developmental scales in mongolism. Journal of Paediatrics 62, 93-7.

KOEGEL and MENTIS (1985) Motivation in Autism: Will They or Won't They? Journal of Child Psychology and Psychiatry 26 (2) 185.

KOSTREWSKI, J. (1974) The dynamics of Intellectual Development in Individuals with Complete and Incomplete Trisonomy of Chromasome Group G in the Karotype. Polish Psychological Bulletin 5 (3) 153-8.

La VECK, B. and La VECK, G. D. (1977) Sex differences among young Down's syndrome children. Journal of Pediatrics (91), 767-9.

LAVECK, B. and BREHM, S. S. (1978) Individual Variability Among Children with DS Mental Retardation 16 (2) 135-7.

LEJEUNE, J., GAUTIER, M. and TURPIN, R. (1959). Etudes des chromosomes somatiques de neuf enfants mongoliens, C. R. Acad. Sci. (Paris) 248 1721-2.

LEWIS, M. and JOHNSON, N. (1971) What's thrown out with the bath water: A baby? Child Development, 42, 1053-1055.

LOVELAND, K. A. (1987). Behaviour of young children with Down Syndrome before the mirror: Finding things reflected. Child Development 58 (4) 928-936.

LUDLOW, J. R. and ALLEN, L. M. (1979) The effect of early intervention and pre-school stimulus on the development of the Down's Syndrome child. Journal of Mental Deficiency Research 23, 29-44.

LUISELLI, J. K. and DONELLON, S. (1980) Use of a Visual Stimulus Fading Procedure to Teach Colour Naming to an Autistic Child. Journal of Behaviour Therapy and Experimental Psychiatry vol 11 73-76.

MacPHERSON, F and BUTTERWORTH, G. (1981) Application of a Piagettian Infant Development Scale to the Assessment of Profoundly Mentally Handicapped Children. Paper presented to Annual Conference, British Psychology Society. (Developmental Psychology Section), Manchester.

MacPHERSON, F. D. (1984). The structure of sensory motor intelligence in severely and profoundly mentally handicapped children. Unpublished PhD thesis, University of Southampton, England.

MacTURK, R., VIETZE, P., McCARTHY, M., McQUISTON, S. and YARROW, L. (1985). The organisation of exploratory behavior in Down's syndrome and non-delayed infants. Child Development, 56, 573-587.

MAISTO, A. A. and GERMAN, M. L. (1986) Reliability, predictive validity and interrelationships of early assessment indices used with developmentally delayed infants and children. Journal of Clinical Child Psychology, 15 (4) 327-332.

MANS, L., CICCHETTI, D. and STROUFE, L. A. 1978: Mirror reactions of Down's Syndrome infants and toddlers: Cognitive underpinnings of self recognition. Child Development, 49, 1247-50.

MARCEL, M. M. and ARMSTRONG, V. (1982). Auditory and visual sequential memory of Down syndrome and non-retarded children. American Journal of Mental Deficiency, 87, (1), 86-95.

MARCEL, M. M., HARVEY, C. F., and COTHRAN, L. P. (1988). An attempt to improve auditiory short term memory in Down's Syndrome. Research in Developmental Disabilities 9 (4), 405-417.

MARCOVITCH, S., GOLDBERG, S., MacGREGOR, D. and LOJKASEK, M. (1986). Patterns of temperament variation in three groups of developmentally delayed preschool children: Mother and father ratings. Development and Behavioral Pediatrics, 7, 247-252.

MARTIN, W. E. and BLUM, A. (1961) Interest Generalisation and Learning in Mentally Normal and Subnormal Children. Journal of Comparative and Physiological Psychology 54 28-32

McCALL, R. B. (1981) Nature-nurture and the two realms of development: A proposed integration with respect to mental development. Child Development, 52, 1-12.

McDADE, H. L. and ADLER, S. (1980). Down Syndrome and short-term memory impairmenta storage or retrieval deficit? American Journal of Mental Deficiency, 84, 561-7.

McIVOR, M. and McGINLEY, P. (1983) Teaching Discrimination Using an Errorless Learning Procedure. Mental Handicap vol 11 September p.110-113.

McLEAN, M., McCORMICK, K., BRUDER, M. B. and BURDG, N. B. (1987) An investigation of the validity and reliability of the Battelle Development Inventory with a population of young children younger than 30 months of age with identified handicapping conditions. Journal of the Division for Early Childhood, 11 (3), 238-246.

McMAHON, B. and PUGH, T. F. (1970). Epidemiology principles and methods. Boston: Little Brown and Co.

MELYN, M. A. and WHITE, D. T. (1973). Mental and developmental milestones of non-institutionalised Down's syndrome children. Pediatrics, 52 (4), 542-5.

MERIGHI, EDISON and ZIGLER, E. Motivation: a developmental perspective. In R.M. Hodapp, J.A. Burack and E. Zigler (Eds.) Issues in the developmental approach to retardation. Cambridge University Press 1990.

MERVIS, C. B. (in press). Early conceptual development of children with Down's syndrome. In D. Cicchetti and M. Beeghly (Eds.), Children with Down syndrome: A developmental perspective. New York: Cambridge University Press.

METCALF, J. A. and STRATFORD, B. (1986). Development of perception and cognitive abilities among non-handicapped children and children with Down Syndrome. Australia and New Zealand Journal of Developmental Disabilities, 12 (1), 65-78.

MIKKELSEN, M. (1977). Down syndrome: cytogenetical epidemiology Hereditas 86 45-50.

MILGRAM, N. A. The Rationale and Irrational in Zigler's motivational Approach to Mental Retardation American Journal of Mental Deficiency 73 527-31.

MILLER, J. F. (in press). Language and communication characteristics of children with Down syndrome. In A. Crocker, S. Pueschel, J. Rynders and C. Tingley (Eds.), Down syndrome: State of the art. Baltimore: Brooks.

MIRANDA, S. B. and FANTZ, R. L. (1973). Visual preferences of Down syndrome and normal infants. Child Development, 44, 555-561.

MIRANDA, S. B. and FANTZ, R. L. (1974). Recognition memory in Down syndrome and normal infants. Child Development, 45, 651-660.

MITCHELL, S. C., KORENEZ, S. B., and BERENDES H. W. (1971). Congenital heart disease in 56, 109 births. Circulation, 43, 323.

MOORE, R and GOLDIAMOND, I. (1964) Errorless Establishment of Visual Discrimination Using Fading Procedures. Journal of the Experimental Analysis of Behaviour vol 7 pp.269-272.

MORGAN, S. B. (1979) Adaptive skills in young Down's Syndrome children: implications for early intervention. Mental Retardation 17 (5), 247-9.

MORSS, J. R. A Comparative Study of the Cognitive Development of the Infant with Down's Syndrome and the Normal Infant. PhD Thesis University of Edinburgh 1979.

MORSS, J. R. Cognitive development in the Down's Syndrome infant: Slow or different? British Journal of Educational Psychology, 53, 40-47.

MORSS, J. R. Early Cognitive Development - Difference or Delay. In B. STRATFORD and D. LANE (Eds) Current Approaches to Down's Syndrome 1985.

MOSS, S. and HOGG, J. (1987). The interaction of manipulative movements in children with Down's Syndrome and their non-handicapped peers. Human Movement Science, 6, 67-99.

MOTTI, F., CICCHETTI, D., and SROUFE, L.A. (1983). From infant affect expression to symbolic play: The coherence of development in Down syndrome children. Child Development, 54, 1168-1175.

MULCAHY, M. T. (1979). Down's Syndrome in Western Australia; mortality and survival Clinical Genetics 16, 103.

MUNDY, P., SIGMAN, S., KASARI, C., YURMIYA, N. (1988). Non-verbal communication skills in Down Syndrome. Child Development, 59 (1), 235-249.

NAGLIERI, J. A. (1981). Extrapolated developmental indices for the Bayley Scales of Infant Development. American Journal of Mental Deficiency, 1981, 85, 548-550.

NOLAN, M., MCCARTNEY, E., MCARTHUR, K. and ROWSON, V. R. (1980). A study of the hearing and receptive vocabulary of the trainees of an adult training centre. Journal of Mental Deficiency Research 24, 271-86.

O'CONNOR, N. and HERMELIN, B. (1961) Visual and Stereognostic Shape Recognition in Normal Children and Mongol and non-Mongol Imbeciles. Journal of Mental Deficiency Research 5, 63-6.

OSTER, J. (1953). Mongolism. Copenhagen: Danish Science Press.

PIAGET, J. (1936). The Origins of Intelligence in Children. London: Routledge and Kegan Paul, 1953 (original French edition 1936).

PIAGET, J. (1937). The Construction of Reality. London: Routledge and Kegan Paul. 1955 (original French edition 1937).

PIAGET, J. (1950). The Psychology of Intelligence. London: Routledge and Kegan Paul.

PIERSE, D. and EUSTACE, P. (1971). Acute keratoconus in Mongols. British Journal of Opthalmology. 55, 50-54.

PIPER, M. C., GOSSELIN, C., GENDRON, M., and MAZER, B. (1986) Developmental profile of young Down's Syndrome infants receiving early intervention. Childs Care, Health and Development, 183-194.

PRUESS, J. B., VADASY, P. F. and FEWELL, R. R. (1987). Language development in children with Down Syndrome: An overview of recent research. Education and Training in Mental Retardation, 22 (1), 44-55.

PUESCHEL, S. M. (Ed) (1984) The Young Child with Down's Syndrome. New York: Human Sciences Press.

PURPURA, D. (1975). Dendritic differentiation in human cerebral cortex: Normal and aberrant developmental patterns. In G. Kretzberg (Ed.), Advances in neurology, Vol 12 (pp. 91-116).

RAMSAY, M. and PIPER, M. C. (1980) A comparison of two developmental scales in evaluating infants with Down's Syndrome. Early Human Development 4 (1), 89-95.

RAUH, H. and RUDINGER, G. (1987) Early development of Down Syndrome children as assessed by the Bayley Scales. In D. LIEPMANN, G. MOHR, R. and SCHWARZER (Eds), Arbeitsberichte des IfP der FU Berlin 4 p 1-34. Berlin Universitatsdruckerei.

RICHARDS, M. R., MERRITT, K. K., SAMULES, M. H. and LANMANN, A. G. (1955). congenital malformations of the cardiovascular system in a series of 6, 053 infants. Pediatrics 15, 12.

RIGRODSKY, S., PRUNTY, F. and GLOVSKY, G. (1961). A study of the incidence, types and associated etiologies of hearing loss in an institutionalised mentally retarded population. Training School Bulletin 58, 30-44.

ROGERS, S. J. (1977). Characteristics of the cognitive development of profoundly retarded children. Child Development, 48, 837-843.

RONDAL, J. A. (1984). Linguistic and pre-linguistic development in moderate and severe mental retardation. In DOBBING, J. et al. (Eds.). Scientific Studies in Mental Retardation. London: Macmillan Press.

RONDAL, J. A. (1988). Language development in Down's Syndrome: a life-span perspective. International Journal of Behavioural Development, 11, 21-36.

ROTHBART, M. and HANSON, M. (1983). Neurobiology of Down's syndrome. Progress in Neurobiology, 21, 199-237.

ROWE R. D. and UCHIDA, I. H. (1960). Congenital cardiac malformations in mongolism. British Heart Journal 22, 331.

RYNDERS, J. E., SPIKER, D, and HORROBIN, J. M. Underestimating the educability of Down's Syndrome children: examination of methodological problems in recent literature. American Journal of Mental Deficiency, 1978 (1982) 440-448.

SCARR, S. (1981) Testing for children: Assessment and the many determinants of intellectual competence. American Psychologist, 36, (10), 1159-1166.

SCARR-SALAPATEK, S. (1976). An evolutionary perspective on infant intelligence: Species patterns and individual variations. In M. LEWIS (Ed.) Origins of Intelligence: Infancy and Early Childhood, (pp 165-198). New York: Plenum Press.

SCHILMOELLER, G. L., SCHILMOELLER, K. J., ETZEL, B. C. and Le BLANC, J. M. (1979) Conditional Discrimination After Errorless and Trial and Error Training. Journal of the Experimental Analysis of Behaviour vol 31 p 405-420.

SCHILMOELLER, K. J. and ETZEL, B. C. (1977) An Experimental Analysis of Criterion Related and Non Criterion Related Cues In 'Errorless' Stimulus Control Procedures. In B. C. ETZEL, J. M. Le BLANC and D. B. BAIER (Eds) New Developments in Behavior Research: Theory, Method and Application. New York: Hillsdale Lawrence Erlbaum Associates.

SCHNELL, R. R. (1984) Psychomotor development. In S. M. PUESCHEL (Ed.), The Young Child with Down Syndrome, p 207-226. New York: Human Sciences Press.

SELIGMAN, M. E. P. Helplessness: On Depression Development and Death. San Francisco: Freeman 1975.

SHARAV, T., and SHLOMO, L. (1988). Stimulation of infants with Down Syndrome: Long term effects. Mental Retardation. 24, 81-86.

SHARE, J., WEBB, A. and KOCH, R. (1961) A preliminary investigation of the early developmental status of mongoloid infants. American Journal of Mental Deficiency 66, 238-241.

SHERMAN, T. W. and WEBSTER, C. D. (1974) The Effects of Stimulus Fading on Acquisition of a Visual Position Discrimination in Autistic Retarded and Normal Children. Journal of Autism and Childhood Schizophrenia 4 no 4 p 301-312.

SIDMAN, M. and STODDARD, L. T. (1966) Programming Perception and Learning for Retarded Children. In N. R. ELLIS (Ed) International Review of Research in Mental Retardation vol 2 pp. 151-208 New York : Academic Press.

SIDMAN, M. and STODDARD, L. T. (1967) the Effectiveness of Fading in Programming a Stimulus Form Discrimination for Retarded Children. Journal of the Experimental Analysis of Behaviour 10 pp. 3-15.

SMITH, J. D. (1981). Down's Syndrome, amniocentesis, and abortion: prevention or elimination. Mental Retardation. 19(1), 8-11.

SPIKER, D. A Descriptive study of mother-child teaching interaction with high and low functioning Down's Syndrome preschoolers. Unpublished doctoral dissertation, University of Minnesota, 1979.

STEDMAN, D. J. and EICHORN, D. H. (1964). A comparison of the growth and development of institutionalised and home-reared mongoloids during infancy and early childhood. American Journal of Mental Deficiency. 69, 391-401.

STERNLICHT, M. and WANDERER, Z. W. (1962). Nature of institutionalised adult mongoloid intelligence . American Journal of Mental Deficiency, 67, 301-2.

TAKASHIMA, S., BECKER, L., ARMSTRONG, D., and CHAN, F. (1981). Abnormal neuronal development in the visual cortex of the human fetus and infant with Down's syndrome: A quantitative and qualitative Golgi study. Brain Research, 225, 1-21.

TERRACE, H. S. (1963a) Discrimination Learning With and Without Errors. Journal of the Experimental Analysis of Behaviour vol 6 1-27.

TERRACE, H. S. (1963b) Errorless Transfer of a discrimination Across Two Continua. Journal of the Experimental Analysis of Behaviour vol 6 p 223-232.

THASE, M. E. (1982). Longevity and mortality in Down's Syndrome. Journal of Mental Deficiency Research, 26, 177.

THOMPSON, J. S & THOMPSON, M. W . Genetics in Medicine (3rd Ed) Toronto, Saunders.

THOMPSON, R., CICCHETTI, D., LAMB, M. and MALKIN, C. (1985). The emotional responses of Down syndrome and normal infants in the strange situation: the organisation of affective behavior in infants. Developmental Psychology, 21, 828-841.

TOUCHETTE, P. E. (1968) The Effects Of Graduated Stimulus Change on the Acquisition of a Simple Discrimination in Severely Retarded Boys. Journal of the Experimental Analysis of Behaviour 1139-48.

VIETZE, P., McCARTHY, M., McQuiston, S., MacTURK, R. and YARROW, L. (1983). Attention and exploratory behaviour in infants with Down syndrome. In T. Field and A. Sostek (Eds.) Infants born at risk: Perceptual and physical processes (pp. 251-268). New York: Grune and Stratton.

WAARDENBURG, P. J. (1932). Das menscliche Auge und seine Erbanlagen. The Hague: Nijhoff.

WACHS, T. D. and SMITHERMAN, C. H. (1985). Infant temperament and subject loss in a habituation procedure. Child Development, 56, 861-867.

WALD, N. J., CUCKLE, H. S., DENSEM, J. W., NANCHAHAL, K., ROYSTON, P., CHARD, T., HADDOW, J. E., KNIGHT, G. J., PALOMAKI, G. E., CANICK, J. A. Maternal serum screening for Down's syndrome in early pregnancy. British Medical Journal. 297, 883-7.

WEINSHILBAUM. R., THOA., N., JOHNSON, D., KOPIN, I. and AXELROD, J. (1971). Proportional release of norepinephrine and dopamine-beta-hydroxylase from the sympathetic nerves. Science, 174, 1349-1351.

WEISZ, J. R., (1982) Learned Helplessness and the Retarded Child. In E. ZIGLER and D. BALLA (eds.) Mental Retardation: The Developmental-Difference Controversy, Hillsdale, N.J. Erlbaum.

WERNER, E. E., and BAYLEY, N. (1966). The reliability of Bayley's revised scale of mental and motor development during the first year of life. Child Development, 37, 39-50.

WERNER, H. (1948). The concept of development from a comparative and organismic point of view. In D. Harriss (Ed.), The concept of development. Minneapolis, MN: University of Minnesota Press.

WISHART, J. G. (1986a) The effects of step-by-step training on cognitive performance in infants with Down's Syndrome. Journal of Mental Deficiency Research, 30, 233-250.

WISHART, J. G. (1986b). Cognitive avoidance patterns in infants and young children with Down's Syndrome. Paper presented at the international Down's Syndrome Congress, Brighton , England, April.

WISHART, J. G. (1987) Performance of 3-5 year old Down's Syndrome and non-handicapped children on Piagetian infant search tasks. American Journal of Mental Deficiency, 92, (2), 169-177.

WISHART, J. G. (1988). Early learning in infants and young children with Down's Syndrome. In NADEL, L. (Ed) The Psychobiology of Down Syndrome. Boston: M. I. T. Press.

WISHART, J. G., (1990) Learning to Learn: The Difficulties Faced by Infants and Young Children with D. S. In W.I. FRASER (Ed.) Key Issues in Mental Retardation. Routledge, London.

WOODWARD, M. (1959) The Behaviour of Idiots Interpreted by Piaget's Theory of Sensori Motor Development. British Journal of Educational Psychology 29 60-71.

WOODWARD, W. M. (1979) Piaget's theory and the study of mental retardation. In ELLIS, N. R. (Ed.) Handbook of Mental Deficiency: Psychological Theory and Research. Hillsdale, NJ. Erlbaum.

WYNN, M. and WYNN, A. (1979) Prevention of Handicap and the health of women. London : Routledge and Kegan-Paul.

YATES, C., SIMPSON, J., MALONEY, A., GORDON, A., and REID, A. (1980). Alzheimer-like cholinergic deficiency in Down's syndrome. Lancet, 2, 979.

ZEAMAN, D. and HOUSE, B. (1963) The Role of Attention in Retardate Discrimination Learning. In N. R. ELLIS (Ed) Handbook of Mental Deficiency. New York McGraw-Hill.

ZIGLER, E. (1966a) Research on Personality Structure in the Retardate. In ELLIS, N. R. International Review of Research in Mental Retardation vol 1 New York Academic Press.

ZIGLER, E. (1969) Developmental versus difference theories of mental retardation and the problem of motivation, American Journal of Mental Deficiency, 73, (4), 536-556.

ZIGLER, E. (1973) The Retarded Child as a Whole Person. In D. K. ROUTH (Ed) The Experimental Psychology of Mental Retardation. London: Crosby Lockwood Staples.

Br. J. educ. Psychol., 57, 265-278, 1987

A COMPARISON OF TWO PROCEDURES FOR TEACHING DISCRIMINATION SKILLS TO DOWN'S SYNDROME AND NON-HANDICAPPED CHILDREN

By LOUISE DUFFY and JENNIFER G. WISHART
(Department of Psychology, University of Edinburgh)

Summary. Two different strategies for teaching discrimination to Down's Syndrome (DS) and non-handicapped children were compared for relative efficiency: trial-and-error and errorless learning. Two types of discrimination tasks were used, shape and nonsense figure tasks. A pre-test was used to match children for pre-existing ability. Errorless learning proved to be the superior training strategy in each group, both during training and in post-tests. DS children responded poorly to trial-and-error training in both absolute and relative terms. Although order of presentation of training conditions had little effect on performance in the non-handicapped group, an interesting differential effect emerged in the DS group: initial trial-and-error training adversely affected subsequent performance in the errorless task while initial errorless experience enhanced subsequent trial-and-error performance. It would appear from these results that errorless learning may be useful as a "primer", increasing motivation to learn in more conventional learning situations.

INTRODUCTION

Down's Syndrome (DS) is the single greatest cause of mental retardation in the UK, accounting for nearly one third of children with severe mental handicap. Huge advances in our understanding of the metabolic and biochemical abnormalities associated with DS have been made recently (see Smith, 1985) but these are unlikely to benefit those children already born with DS. Prenatal screening has made some inroads into incidence rates but medical and economic factors at present limit the use of screening to women already known to be at risk, that is, women over 35 or women who have already given birth to a child with DS; even within this high risk group, it would appear that a significant proportion are either not offered screening or do not take up that offer (Walker and Howard, 1986). This at-risk group accounts in any case for only one third of all DS births. Prevention of DS by use of pre-natal screening techniques is still, therefore, a distant prospect. Recent studies, in fact, show little decline in overall incidence rates, with figures suggesting an increase in infants born to mothers in younger age groups (Abroms and Bennett, 1983; Stratford and Steele, 1985). Prevalence, additionally, has increased four-fold within the last generation as a result of advances in medical technology, particularly in paediatric cardiology (Kirman, 1983).

As long as prevalence and incidence figures remain high, it would be unwise to abandon psychological approaches to the study of DS on the grounds that future discoveries in the medical field may ameliorate — or even eliminate — the condition. The anomalies associated with DS are expressed behaviourally. Psychology can, therefore, contribute much to attempts at facilitative intervention. It seems essential that we continue to try to learn more about the exact nature of the mental handicap found in children with DS. It seems particularly important that we attempt to find some way to counteract the progressive decline in rate of mental development generally found with increasing age (Gibson, 1978).

Recent years have seen a significant impact on development as a result of DS children being brought up in the parental home rather than in an institutional setting (Centerwall and Centerwall, 1960; Carr, 1975, 1985). In normal children, variations in ability, particularly in IQ, are often associated with factors such as parental IQ and social class. This relationship does not generally hold true in DS (Gibson, 1978; Carr, 1985), although, interestingly, exceptions to this trend have recently been reported (Sharev *et al.*, 1985; Cunningham, 1986). Given the wide variation in learning ability found in children with DS and this absence of any clear, straightforward relationship with environmental factors usually closely associated with developmental outcome, it would seem that other, less obvious factors must be influencing development in DS. If these could be identified, appropriate fine-tailoring of the environment to the particular needs and skills of the DS child might well improve on presently achieved levels of development. It is important, however, to be realistic about what may be achieved: the genetic component in DS must inevitably set some upper limit on improvement in developmental outcome. Nonetheless, there seems reason to believe that with appropriately sensitive teaching methods, more children with DS could be encouraged to develop to their full potential.

Two distinct psychological approaches to research into learning difficulties in mentally handicapped children dominate the literature, one based in cognitive theories of mental handicap, the other emphasising motivational deficits. The cognitive approach emphasises deficits in cognitive functioning as the source of learning difficulties: development in the mentally handicapped is generally characterised as identical in nature to cognitive development in normal children, with only rate and end-point of development differentiating the two populations (Illingworth, 1980). In contrast, motivational theorists emphasise the role of motivational deficiencies, often classifying the mentally handicapped as "failure avoiders" (Cromwell, 1967): more motivated to avoid failure than to achieve success, the mentally handicapped are characterised as avoiding learning situations and responding poorly to traditional trial-and-error teaching methods (for reviews, see Zigler and Balla, 1982). The motivational and cognitive approaches outlined rarely overlap in any investigation of learning processes in mentally handicapped children. It is possible, however, that the cognitive deficits seen in the DS child may to some extent *result* from motivational problems: frequent early experience of failure may erode motivation to learn, thereby contributing directly to subsequent deficits in functioning.

It is, of course, impossible to examine motivation or cognition in isolation from each other. Terms such as "cognitive style" acknowledge this. However, different teaching strategies place differing emphases on these two performance factors. The study reported here sought to compare two different methods for teaching discrimination to young children with DS: trial-and-error learning and errorless learning.

Errorless learning is the most widely-used strategy for teaching new skills to the mentally handicapped. It has proved highly effective in teaching adolescents practical discrimination skills, skills previously unlearned when conventional methods were used (Cullen, 1976; McIvor and McGinley, 1983; Adams, 1984). To some degree, its adoption implies recognition of the possibility that motivational factors may influence cognitive outcome, but generally its use stems from the more pessimistic belief that the experience of erring in a learning situation will always be counter-productive in the mentally handicapped, adding to already existing cognitive problems.

Despite the widespread use of errorless teaching strategies with all ages of handicapped children, few published studies have directly compared its efficiency with more traditional teaching methods. The most interesting papers generally report single, older case studies and often only anecdotal evidence of failure of other teaching methods is given. Many learning theorists — in particular, Piaget (1936, 1937, 1950) — would argue that erring is an essential element in the learning process and that error has as much to contribute to processes of cognitive development as success. Evidence of poor learning ability using traditional trial-and-error methods is not in itself evidence of inability to profit cognitively from mistakes: there has been little direct evidence to support the view that the mentally handicapped cannot also learn from their own mistakes. It is possible, then, that adoption of an exclusively errorless approach throughout development may in some way be perverting or undermining the normal course of early learning.

In the study presented here, the relative efficiency of errorless and trial-and-error methods in training discrimination skills in both DS and non-handicapped children was investigated. Two sets of tasks were used: shape discrimination and nonsense figure discrimination tasks. The use of nonsense figure tasks served as a control for any positive or negative effects that differential prior learning experience with the target concept might have on test performance; balancing of order of presentation of the two training conditions allowed examination of any within-session effects of exposure to the two kinds of learning situation.

Three main hypotheses were tested:

(1) That errorless learning would have a greater enhancing effect on performance than trial-and-error learning and that that effect would be greater in DS than in non-handicapped children.
(2) That initial experience of errorless learning would enhance performance on a subsequently presented trial-and-error task and that this enhancement would be greater in DS children.
(3) That initial experience of trial-and-error learning would adversely affect performance on a subsequently presented errorless task and that this effect would be present in DS children only.

METHOD

Sample matching

Direct mental age (MA) matching of handicapped and non-handicapped children was avoided for a number of reasons, both practical and theoretical. The validity of MA matching in handicap studies has frequently been questioned (Clarke and Clarke, 1975; Woodward, 1979; Wishart, 1986a). The MA composite is arrived at by simple addition of scores on a number of test items. This means that two children with widely differing ability profiles can achieve identical MAs. Even in the normal population, some "matches" must for some purposes be inaccurate. Even more vulnerable to criticism are MA matches of mentally-handicapped children to much younger non-handicapped children; the greater the chronological age gap, the more widely-differing the individual learning histories of the MA-matched children and the more inexact — and less meaningful — any such "match" is likely to be (Hogg and Moss, 1983).

More importantly, perhaps, acceptance of psychometrically-based matching procedures implies acceptance of one particular theory of mental development in the handicapped, the "slow development" theory. This theory maintains that processes of cognitive development in the handicapped and non-handicapped are identical in nature, with only rate and end-point of development differing (see above). Recent research suggests, however, that learning in DS children may be radically different from that seen in normal children, with important qualitative as well as quantitative differences in learning processes existing in the two populations (Morss, 1983, 1985; Wishart, 1986b). Use of MA-matching involves accepting, moreover, that motivational factors do not differentially influence performance in the two groups, that performance and competence are similarly linked in both populations. Again, there is evidence to suggest that this may not be the case (Byck, 1969; Balla and Zigler, 1979; Shaw, 1986; Wishart, 1987).

Given the particular nature and aims of the study to be reported here, it seemed important to include neither of these theoretical assumptions in the experimental design. Children were therefore matched, not on MA, but on the basis of ability to pass a pre-test based on the discrimination tasks to be used in the experiment itself (see below).

Down's Syndrome group. Thirteen children were selected from two Edinburgh special schools. Three were excluded on the basis of the pre-test (see below) and two had to be dropped because of inability to concentrate for long enough to take part effectively in the experiment. This left eight DS children, five males and three females, who completed the experiment (mean age: 7 years 9 months; SD: 14.6 months).

Non-handicapped group. Thirteen children were selected from two Edinburgh nursery schools. Four were already able to pass the pre-test and one further child had to be excluded later in the experiment because of attention difficulties. Mean age of the remaining eight non-handicapped children was 2 years 6 months (SD: 4 months). There were four males and four females.

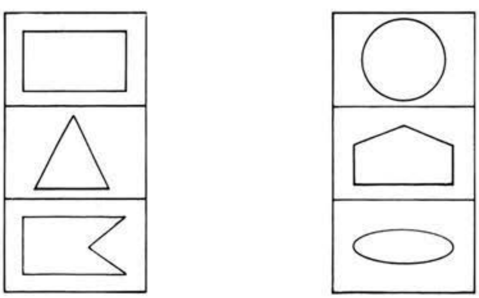
Procedure

Testing of all children took place in a small room, as free from distractions as could be arranged within the school settings. Child and experimenter sat opposite each other at a small table.

Selection pre-test. One of the two target stimuli (a rectangle or oval) was presented together with two alternative stimuli: one "orthodox" shape (circle, square, triangle or parallelogram) and one "unorthodox" shape, introduced to limit the possibility that the oval or rectangle would be "correctly" identified by a process of elimination of more familiar, known shapes (see Figure 1).

FIGURE 1

TEST CARDS USED IN SUBJECT SELECTION PRE-TEST



All shapes were of different colours and approximately 7 cms on their longest dimension. Each was centred on a white card, 12 cm × 8 cm. Position of shape to be identified was randomised over trials. Children were told: "I am going to show you three cards with shapes on them. When I say the name of a shape, I want you to point to the card with that shape on it."

A minimum of 28 trials were given, a minimum of six of each target stimulus and four of each orthodox shape. Further trials were given as necessary until it could be established whether correct responses were due to random guessing or true ability to discriminate the shape in question. Responses were not differentially reinforced in this pre-test. Children were selected on the basis that they showed an ability to discriminate some shapes but were not yet able to discriminate rectangles and ovals.

All children chosen on the basis of the selection pre-test went on to be presented with the pre-tests, training and post-tests of the shape discrimination tasks, followed by the training and post-tests of the nonsense figure discrimination tasks. (No pre-test was required for either of the nonsense figure tasks since no prior knowledge could safely be assumed). Both groups of children were divided into two subgroups and order of presentation of errorless and trial-and-error training within each discrimination task balanced over subgroups and for each child.

(A) Shape discrimination tasks

Trial-and-error training used an oval as the target stimulus, errorless training a rectangle.

Pre- and post-tests. Seven sets of cards, identical in format to those used in the selection pre-test, were used. (Initial use of 10 trials in both the pre- and post-test resulted in an unacceptably long procedure.) The target stimulus varied in size, colour and proportion over trials in order to represent the concept of rectangle or oval in its most general form. Position of target stimulus was randomised over trials. Instructions to children were as in the selection pre-test.

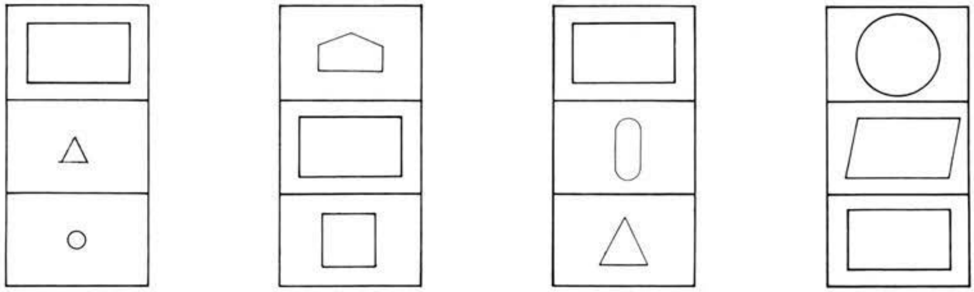
(i) Trial-and-error training (oval). As in the pre-test, an oval was presented together with a common and an unorthodox shape. 15 trial sets of three cards were used. Colour but not size was varied over trials. Order of presentation of the 15 cards was randomised.

Children were told: "I am going to show you three cards with shapes on them just as I did the last time and I want you to point to the card with the shape I ask for. This time I will tell you if you are right." Correct responses were verbally praised. Incorrect responses were negatively reinforced; children were told: "No, that is not right. Try again the next time."

(ii) Errorless training (rectangle) 15 trial sets of three cards were used. The trials were presented in an order by which two alternative stimuli to the rectangle were gradually "faded in", increasing in size while varying in shape over trials (see Figure 2).

FIGURE 2

EXAMPLES OF ERRORLESS LEARNING TEST CARDS
(SHAPE DISCRIMINATION — RECTANGLE)



In the first two sets, the target stimulus was presented with two blank cards. In these and all subsequent trials, position of the target stimulus was randomised over trials, as was colour, the latter a precaution against misidentification of this as a relevant attribute of the target stimulus. Trials 3 and 4 consisted of a rectangle with two similarly coloured but much smaller (0.5 cm) alternative shapes. The dimensions of the alternative shapes were gradually increased over trials 5-13. The final two trials consisted of three red shapes, a rectangle and two alternatives of equivalent dimensions.

Verbal instructions to the children were identical to those given in the trial-and-error training. Verbal praise was given each time the child made a correct response. Errors were not commented on but, rather than proceeding with the next trial, the previous trial was re-presented, this procedure being repeated as necessary, until the child had shown mastery of that particular step in the training sequence.

(B) Nonsense figure discrimination tasks

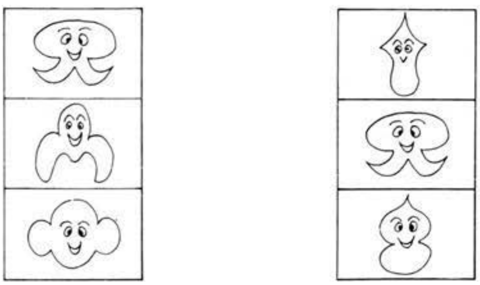
Trial-and-error training used "nims" as the target stimulus, errorless training used "wugs" (see below).

Pre-tests. No pre-tests were presented since no prior knowledge of the concept of "nims" or "wugs" could be assumed.

(i) Trial-and-error training ("nims"): 15 trial sets of three cards were presented, each showing a "nim" and two other nonsense figures of equivalent size and colour (see Figure 3).

FIGURE 3

EXAMPLES OF TRIAL-AND-ERROR LEARNING TEST CARDS
(NONSENSE FIGURE DISCRIMINATION — NIMS)



Children were shown a drawing of a Mr Plimp — a "nim"-like figure — and told that he had some friends called "nims" in the pack of cards on the table. The aim of the game was to help Mr Plimp to find all the "nims". They were then told: "I am going to show you three cards with little men on them. I want you to point to the card you think has Mr Nim on it. Correct responses were verbally praised on behalf of Mr Plimp and when an incorrect response was made, children were told to try again next time.

(ii) Errorless training ("wugs"): 15 trial sets of three cards were used. Alternative stimuli to the "wug" were faded in over trials in exactly the same way as in the errorless rectangle training.

Post-tests. Post-test trials consisted of seven test card sets, each with either a "nim" or a "wug" and two other similarly sized nonsense figures.

RESULTS

All responses in the pre-tests, training and post-tests were used in the analysis. Scores in the pre- and post-tests were expressed in terms of the number of correct responses made. Since the number of trials presented in the errorless procedure was child-determined, scores in the errorless and trial-and-error training periods were expressed as percentages (correct responses/total responses).

Shape discrimination

Pre-test trials. Pre-test scores on both shape tasks were compared in the DS and non-handicapped groups. No significant differences were found, thereby validating the procedure adopted for sample selection and matching (oval (trial-and-error): $t = 1.078$, df (14), NS; rectangle (errorless): $t = 1.56$, df (14), NS). There were also no differences within groups in pre-test performance on the two discrimination tasks (DS: $t = 0.74$, df (7), NS; non-handicapped: $t = 0.42$, df (7), NS).

Training trials. Table 1 shows the percentage of correct responses during trial-and-error and errorless shape training trials for both the DS and the non-handicapped children. Performance in the two training conditions differed significantly in the combined groups ($t = 3.7$, df (15), $P < 0.005$) and also in both groups taken separately (DS: $t = 11.4$, df (7), $P < 0{\cdot}0005$); non-handicapped: $t = 4.8$, df (7), $P < 0.005$), with errorless training scores exceeding trial-and-error training scores in all three cases.

T-tests comparing the trial-and-error training scores of children initially trained with errorless learning and children with no prior training revealed a significant

TABLE 1

PERCENTAGE OF CORRECT RESPONSES DURING SHAPE TRAINING TRIALS

	Errorless Training (rectangle)	Trial-and-error Training (oval)
Initial training	*DS subjects*	
	1) 100	60
Errorless	2) 100	46
	3) 100	66
	4) 100	33
	5) 80	40
Trial-and-error	6) 85	46
	7) 100	33
	8) 100	40
	Non-handicapped subjects	
	1) 82	60
Errorless	2) 94	60
	3) 89	46
	4) 100	100
	5) 100	66
Trial-and-error	6) 94	33
	7) 100	26
	8) 94	46

difference in favour of those who had experienced the errorless training prior to presentation of the trial-and-error training ($t = 2.15$, df (14), $P < 0.025$). A similar, if weaker, effect existed in both groups when analysed separately (DS: $t = 1.47$, df (6), $P < 0.10$; non-handicapped: $t = 1.63$, df (6), $P < 0.10$). No effect of order of presentation of training conditions on errorless scores was found although there seemed to be a trend, in the DS group only, in favour of better performance in errorless training in children with no prior experience of trial-and-error shape learning ($t = 1.70$, df (6), $P < 0.10$).

Post-test trials. Table 2 shows the differential effects of the two training procedures on subsequent performance. Improvement in performance was calculated by comparing pre- and post-test scores. For the combined groups, performance improved more following errorless training ($t = 1.78$, (df 15), $P < 0.05$). This effect did not reach significance in either group, however, although it came closest to doing so in the non-handicapped group ($t = 1.76$, df (7), $P < 0.10$).

There was no overall effect of order of training procedures on post-test improvement. Only the trial-and-error scores in the DS group varied significantly with order of presentation, with prior errorless experience enhancing subsequent trial-and-error performance ($t = 1.96$, df (6), $P < 0.05$). Improvement in performance in the two non-handicapped sub-groups was roughly equivalent in both training conditions, irrespective of their order of presentation.

TABLE 2

SHAPE DISCRIMINATION — EFFECTS OF ERRORLESS AND TRIAL-AND-ERROR TRAINING ON SUBSEQUENT PERFORMANCE: PRE-/POST-TEST SCORE DIFFERENCES

	Errorless Training Scores (rectangle)			Trial-and-Error Training Scores (oval)		
	Pre-	Post-	Diff.	Pre-	Post-	Diff.
Initial training			*DS subjects*			
	1) 0	5	5	2	5	3
Errorless	2) 0	7	7	5	7	2
	3) 4	7	3	2	7	5
	4) 2	4	2	1	6	5
	5) 1	2	1	1	2	1
Trial-and-error	6) 0	5	5	2	5	3
	7) 1	5	4	2	1	-1
	8) 3	6	3	1	4	3
			Non-handicapped subjects			
	1) 1	5	4	4	7	3
Errorless	2) 3	6	3	2	6	4
	3) 2	6	4	2	3	1
	4) 6	7	1	7	7	0
	5) 4	7	3	2	7	5
Trial-and-error	6) 2	6	4	1	5	4
	7) 2	7	5	2	2	0
	8) 1	6	5	3	6	3

Nonsense figure discrimination

Training trials. Errorless and trial-and-error training scores are set out in Table 3. Errorless training scores reliably exceeded trial-and-error training scores, both overall and in each group, although this difference in scores was less pronounced in the non-handicapped group (overall: $t = 6.15$, df (15), $P < 0.0005$; DS: $t = 5.96$, df (7), $P < 0.0005$; non-handicapped: $t = 3.31$, df (7), $P < 0.01$).

TABLE 3

PERCENTAGE OF CORRECT RESPONSES DURING NONSENSE FIGURE TRAINING TRIALS

	Errorless training (Wug)	Trial-and-error training (Nim)
Initial training	*DS subjects*	
	1) 94	40
Trial-and-error	2) 100	26
	3) 82	13
	4) 94	80
	5) 100	33
Errorless	6) 94	46
	7) 100	40
	8) 100	86
	Non-handicapped subjects	
	1) 100	33
Trial-and-error	2) 100	100
	3) 94	66
	4) 100	73
	5) 100	33
Errorless	6) 94	80
	7) 94	53
	8) 100	73

Comparison of trial-and-error training scores of children initially given the errorless task and children with no prior training showed no significant differences, either for the combined groups or for either group taken separately. Nor was there any effect of training order on errorless training scores, although, as in the shape task, a trend in favour of the DS subgroup given this training condition first, i.e. with no prior trial-and-error experience, was present ($t = -1.48$, df (6), $P < 0.10$).

Post-test trials. Since there were no pre-test scores, the differential effect of the two training strategies on nonsense figure discrimination was evaluated by comparison of post-test scores in the two training groups (see Table 4). Errorless learning proved superior in all comparisons, with the effect again more pronounced in the DS group than in the non-handicapped group (overall results: $t = 3.6$, df (15), $P < 0.005$; DS group: $t = 3.2$, df (7), $P < 0.01$; non-handicapped group: $t = 2.18$, df (7), $P < 0.05$). Although a trend in favour of initial errorless learning was evident in the post-test scores of both the non-handicapped and DS groups, the differences failed to reach significance, either overall or for either group.

TABLE 4

DIFFERENCE BETWEEN NONSENSE FIGURE POST-TEST SCORES AFTER ERRORLESS AND TRIAL-AND-ERROR TRAINING

	Errorless training	Trial-and-error training	Diff.
Initial training		*DS subjects*	
	1) 5	3	2
Trial-and-error	2) 7	0	7
	3) 7	5	2
	4) 6	3	3
	5) 7	0	7
Errorless	6) 6	0	6
	7) 7	7	0
	8) 7	7	0
		Non-handicapped subjects	
	1) 7	7	0
Trial-and-error	2) 7	7	0
	3) 6	3	3
	4) 7	7	0
	5) 6	3	3
Errorless	6) 6	5	1
	7) 6	4	2
	8) 7	7	0

DISCUSSION

This study aimed to investigate and compare the efficiency of errorless and trial-and-error methods in teaching discrimination skills to normal and mentally handicapped children. As would have been expected, the non-handicapped group performed better than the handicapped group under trial-and-error training conditions. This kind of training most closely approximates the conditions encountered in everyday, natural learning situations, situations in which the non-handicapped children, by definition, show superior learning skills. It was hoped, however, that the DS children might benefit more than the non-handicapped children from errorless training, in relative if not in absolute terms. Surprisingly perhaps, both groups of children appeared to benefit equally from errorless training, both during training and in the subsequently administered post-tests. While better training scores would have been expected almost by definition with errorless training, its enhancing effect on DS post-test scores is encouraging, particularly given that only one short training session was used. With larger groups and repeated training sessions, it would seem reasonable to hope for even stronger carry-over effects.

From the above, it is clear that the two groups responded very similarly to errorless training but very differently to trial-and-error training. While the differential effect in favour of errorless training existed in both groups, it was more pronounced in the DS group, in both training and post-test score comparisons. Again, to some degee, this was to be expected. For whatever reasons — ability or motivational differences — the trial-and-error training scores of the non-handicapped children would be predicted to be — and were — greater than those of the handicapped children, thereby narrowing the difference between training scores achieved under the two conditions. No such simple explanation can account for the differential effect on post-test scores. After trial-and-error training, post-test scores of the DS children were markedly inferior to those achieved by the non-handicapped children; after errorless training, however, post-test scores on both discrimination tasks matched those achieved by the non-handicapped children. The enhancement effect was particularly noticeable in the nonsense figure results, where errorless training doubled the post-test success rate of the DS children.

Why should the differential effect of the two training conditions be so marked in the handicapped group? It was suggested in the introduction that motivational factors may adversely affect the expression of competence in performance in DS. If motivational factors *do* influence performance in handicapped children over and above cognitive limitations on their performance, it might be expected that initial successful experience with errorless learning, by increasing motivation to learn, might have a priming effect on subsequent performance; initial trial-and-error learning experience might perhaps have the opposite effect, adversely affecting performance and reducing motivation to perform to full potential on the second discrimination task, even with errorless training. No such within-session effect would be predicted in the non-handicapped scores. Given the superior ability and better balanced history of failure and success in this group, there is no reason to expect that their motivation to learn would be either reduced by trial-and-error experience or increased to any significant degree by errorless experience. If anything, experience of *either* form of training, it might be predicted, would beneficially affect subsequent performance, simply by increasing familiarity with the type of task, a straightforward practice effect.

Some limited support for the motivational hypothesis can be found by comparing training and post-test results in the two training subgroups of each group of children. Order of presentation of the two training conditions did appear to influence scores although the effect was not a strong one and, when present, was sometimes to be found in both groups of children. In both shape and nonsense figure discrimination tasks, for example, training scores favoured both of the subgroups given errorless training first, i.e., the children with no prior trial-and-error experience of the target concept; these trends in favour of errorless training were not, however, statistically significant in either group of children on either task. For DS children only, a statistically significant effect of training order was present in scores on the shape task; on the nonsense figure task again only a trend in favour of prior errorless training was evident. While these results are not very strong, they *are* consistent: in all cases where a trend was evident that trend was, without exception, in favour of children given errorless training first; in all cases, the effect was either greater or present only in the DS children.

Some further support for the suggestion that prior, unsuccessful trial-and-error experience of a learning task can negatively affect training outcomes can be found by examining the differential enhancement effects of errorless training in the two discrimination tasks. In the nonsense figure tasks, tasks on which neither group could have had any prior learning experience, the enhancement effect of prior errorless training was clearly demonstrated, with the DS group benefiting most from prior error-free experience. A far weaker enhancement effect of errorless training procedures was found in the shape discrimination task results, both in the training and post-test scores. All children were likely to have had some prior experience of shape teaching, either formal or informal. Given that neither group of children were proficient yet in shape discrimination, any such prior learning experience must, of necessity, have been of a trial-and-error nature.

Although showing similar levels of pre-test ability, learning histories prior to training must, however, inevitably have differed in the two groups of children, even if only in terms of length. Given the difference in general ability levels of the two groups, it seems reasonable to assume that in their lengthier pre-history, DS children would also have been exposed to a higher absolute and relative rate of failure than the non-handicapped children. This prior negative experience could perhaps account for the poorer training outcome with these discrimination tasks; prior experience of failure was actually being added to in the testing situation, further lowering expectation of success, even on the errorless learning task.

Overall, then, the results suggest that errorless learning techniques may be of value both in themselves and by virtue of the priming effect they appear to have on subsequently presented trial-and-error learning tasks. The fact that trial-and-error training appeared to depress DS performance, even on an errorless task, underlines the importance of ensuring that expectations in the handicapped do not lie in the direction of failure. It is easy to see how constant failure due to intrinsic deficits in functioning could negatively influence general expectations of success in any learning situation, resulting ultimately in a form of learned helplessness (Seligman, 1975).

Previous research suggests caution, however, in the use of errorless training. Dweck (1975), for instance, has questioned whether the most effective way of dealing with the poor response of mentally-handicapped children to failure is to eliminate the possibility of failure: teaching the child to respond positively to failure would seem in the long run to be more productive. Errorless learning is, after all, a very artificial teaching strategy, bearing little relation to real-life learning situations. Over-use could lead to an unhealthy and counterproductive dependence on this sort of learning support. Research also suggests that, while effective in teaching specific discriminations, errorless learning does not generalise readily to other tasks or even to post-tests of the same task (Gollin and Savoy, 1968; Etzel *et al.*, 1981). The principles of errorless learning are, moreover, difficult to apply to higher level, more abstract learning tasks. Errorless learning may, however, play a useful role in the progression towards these more advanced forms of learning. Instead of imposing a difficult learning task on an apprehensive learner, errorless learning, by changing the success/failure rate that would normally be experienced, can perhaps be used to teach the child that he/she *can* learn and that learning can be easy, hopefully thereby raising motivation to put this newly realised skill to further use.

Some anecdotal evidence from the study reported here can be offered in support of the contention that it is important to avoid the establishment of a "failure set" being generalised to areas of comparative or untested strength (see also Duffy, 1986). Patterns of performance in at least two of the DS children — children 2 and 3, both 10-year-olds — suggested that the low scores achieved in the selection pre-test did not accurately reflect these children's current level of shape discrimination ability. When questioned at the end of the experiment, child 2 in fact readily admitted that he had deliberately underperformed in the earlier parts of the experiment (although this had not been obvious in any way to the experimenter at the time).

Similarly, two other DS children (also 10-year-olds) who were eventually excluded from the experiment, produced overly consistent *in*correct responses in the selection pre-test; when offered tangible reinforcement for correct responses both changed to near perfect ability to discriminate the shapes in question (see also Byck, 1969; Shaw, 1986). According to the teacher, in one case at least, this behaviour was likely to have been a deliberate tactic produced to avert the possibility of being presented with a more difficult task; this particular child was always reluctant to admit to being able to perform certain tasks even when well within his abilities, apparently suspicious that good performance on an "easy" task would lead to being tested on more difficult tasks, where he was likely to experience failure.

This reluctance to perform to full potential seemed to be a characteristic of only the older DS children tested. A 6-year-old DS boy who had to be excluded from the study made no attempt to hide his capabilities in the selection pre-test. The older DS children, by contrast, rather than allowing themselves to be placed in a situation over which they might have little control, seemed to be imposing their own control over the situation from the start, their poor performance very much a case of "won't do" rather than "can't do" (Koegel and Mentis, 1985). Whether this behaviour is characteristic only of DS or is a product of age in combination with mental retardation remains to be investigated. Two further mentally-handicapped but non-DS 10-year-olds were run through the experimental procedures; although differing widely in levels of discrimination ability, neither showed any tendency to underperform. A further larger-scale investigation of these aspects of DS performance would seem therefore to be worthwhile.

ACKNOWLEDGMENTS.—We would like to acknowledge, with many thanks, the help of the staff of the Special Needs schools and of the Psychology Department and Colinton nursery schools; their cooperation — and patience — was much appreciated. We are also grateful to Lothian Region Department of Education for permission to carry out the research in their schools. Thanks are also due to Liz Deans of Queen Margaret College for very helpful advice. We would especially like to thank all of the children who participated in this study. This research was supported by MRC Project Grant No. G8314998N.

Correspondence and requests for reprints should be addressed to Dr. Jennifer G. Wishart, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, Scotland.

REFERENCES

ABROMS, K. I., and BENNETT, J. W. (1983). Current findings in Down Syndrome. *Except. Child.*, 49, 449-450.

ADAMS, C. S. (1984). Learning by your mistakes: an experimental investigation. Unpublished PhD thesis, University of Manchester.

BALLA, D., and ZIGLER, E. (1979). Personality development in retarded persons. In ELLIS, N. R. (Ed.), *Handbook of Mental Deficiency* (2nd ed.). Hillsdale, NJ: Erlbaum.

BYCK, M. (1969). Cognitive differences amongst different diagnostic groups of retardates. *Am. J. ment. Defic.*, 73, 97-101.

CARR, J. (1975). *Young Children with Down's Syndrome*. London: Butterworth.

CARR, J. (1985). The development of intelligence. In LANE, D., and STRATFORD, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

CENTERWALL, D. J., and CENTERWALL, W. R. (1960). A study of children with mongolism reared in the home compared to those reared away from home. *Pediatrics*, 25, 678-685.

CLARKE, A. D. B., and CLARKE, A. M. (1975). *Recent Advances in the Study of Subnormality*. London: Mind.

CROMWELL, R. L. (1967). Success-failure reactions in mentally retarded children. In ZUBLIN, J., and JERVIS, G. A. (Eds.), *The Psychopathology of Mental Deficiency*. New York: Grune and Stratton.

CULLEN, C. (1976). Errorless learning with the retarded. *Nursing Times*, 72, 45-47.

CUNNINGHAM, C. C. (1986). Patterns of development in Down's Syndrome. Paper read at International Down's Syndrome Congress, Brighton, April. (Proceedings to be published by Wiley, London, 1988.)

DUFFY, L. (1986). Comparison of two procedures for teaching discrimination to Down's Syndrome and normal children. Unpublished MA Honours thesis, Department of Psychology, University of Edinburgh.

DWECK, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *J. Person. soc. Psychol.*, 31, 674-685.

ETZEL, B. C., LeBLANC, J. M., SCHILMOELLER, K. J., and STELLA, M. E. (1981). Stimulus control procedures in the education of young children. In BIJOU, S. W., and RUIZ, R. (Eds.), *Behaviour Modification: Contributions to Education*. Hillsdale, NJ: Erlbaum.

GIBSON, D. (1978). *Down's Syndrome: The Psychology of Mongolism*. Cambridge, Cambridge University Press.
GOLLIN, E. S., and SAVOY, P. (1968). Feeding procedures and conditional discrimination in children. *J. exp. Analysis Behav.*, 11, 443-451.
HOGG, J., and MOSS, S. C. (1983). Prehensile development in Down's Syndrome and non-handicapped preschool children. *Br. dev. Psychol.*, 1, 189-204.
ILLINGWORTH, R. S. (1980). *The Development of the Infant and Young Child: Abnormal and Normal.* (7th ed.). Edinburgh: Churchill Livingstone.
KIRMAN, B. (1983). A current appraisal of Down's Syndrome. *Br. J. hospit. Med.*, 29, 293-296.
KOEGEL, R. L., and MENTIS, M. (1985). Motivation in childhood autism: can they or won't they? *J. child Psychol. Psychiat.*, 26, 185-191.
MCIVOR, M., and MCGINLEY, P. (1983). Teaching discrimination using an errorless learning procedure. *Ment. Handicap*, 11, 110-113.
MORSS, J. R. (1983). Cognitive development in the Down's Syndrome infant: slow or different? *Br. J. educ. Psychol.*, 53, 40-47.
MORSS, J. R. (1985). Early cognitive development: difference or delay? In LANE, D., and STRATFORD, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.
PIAGET, J. (1936). *The Origins of Intelligence in Children*. London: Routledge and Kegan Paul, 1953 (original French edition, 1936).
PIAGET, J. (1937). *The Construction of Reality*. London: Routledge and Kegan Paul, 1955 (original French edition, 1937).
PIAGET, J. (1950). *The Psychology of Intelligence*. London: Routledge and Kegan Paul.
SELIGMAN, M. E. P. (1975). *Helplessness*. San Francisco: W. H. Freeman.
SHAREV, T., COLLINS, R., and SCHLOMO, L. (1985). Effect of maternal education on prognosis of development in children with Down's Syndrome. *Paediatrics*, 76, 387-391.
SHAW, J. (1986). An investigation of some of the motivational factors affecting the performance of Down's Syndrome children. Unpublished MA Honours thesis, Department of Psychology, University of Edinburgh.
SMITH, G. E. (Ed.) (1985). Molecular structure of the Number 21 Chromosome and Down Syndrome. *Ann. NY Acad. Sci.*, 450 (entire vol.).
STRATFORD, B., and STEELE, J. (1985). Incidence and prevalence of Down's Syndrome — a discussion and report. *J. ment. Defic. Res.*, 29, 95-107.
WALKER, S., and HOWARD, P. J. (1986). Cytogenetic prenatal diagnosis and its relative effectiveness in the Mersey Region and North Wales. *Prenatal Diagnosis*, 6, 13-23.
WISHART, J. G. (1986a). The effects of step-by-step training on cognitive performance in infants with Down's Syndrome. *J. ment. Defic. Res.*, 30, 233-250.
WISHART, J. G. (1986b). Cognitive avoidance patterns in infants and young children with Down's Syndrome. Paper presented at the International Down's Syndrome Congress, Brighton, England, April. (Proceedings to be published by Wiley, London, 1988.)
WISHART, J. G. (1987). Performance of 3-5 year old Down's Syndrome and non-handicapped children on Piagetian infant search tasks. *Am. J. ment. Defic.*, 92 (2), 169-177.
WOODWARD, W. M. (1979). Piaget's theory and the study of mental retardation. In ELLIS, N. R. (Ed.), *Handbook of Mental Deficiency: Psychological Theory and Research*. Hillsdale, NJ: Erlbaum.
ZIGLER, E., and BALLA, D. (Eds.) (1982) *Mental Retardation: The Developmental-Difference Controversy*. Hillsdale, NJ: Erlbaum.

(Manuscript received 16th July, 1986)

Br. J. educ. Psychol., 60, 10-22, 1990

INSTABILITY OF PERFORMANCE ON COGNITIVE TESTS IN INFANTS AND YOUNG CHILDREN WITH DOWN'S SYNDROME

BY JENNIFER G. WISHART AND LOUISE DUFFY
(Edinburgh Centre for Research in Child Development)

SUMMARY. Stability of cognitive performance in assessment situations was investigated in a group of 18 children with Down's Syndrome (DS) aged 6 months to 4 years. Two widely used tests of early cognitive development were presented: the Mental Scale of the Bayley Scales of Infant Development and a series of Piagetian object concept tasks. Both sets of tasks were presented in the same testing session, with testing repeated one to two weeks later. Both qualitative and quantitative aspects of performance were recorded. Even when overall scores were similar, patterns of pass/fails were very different in the two testing sessions. Successes showed poor stability over sessions with fails often occurring by default, the result of a refusal to engage fully in a particular task. This pattern of results suggests: (1) that single-session testing may not adequately assess cognitive status in DS children; (2) that the relationship between performance and competence may be unstable in such children, and (3) that "slow development" theories may not adequately describe cognitive development in young children with DS.

INTRODUCTION

WHILE considerable number of children with Down's Syndrome (DS) may also have motor or sensory handicaps, the major handicap is clearly cognitive. DS is the most common form of mental handicap in the UK and world-wide, accounting for approximately one-third of all children with severe mental handicap. DS provides researchers with a unique opportunity to investigate developmental processes in handicap: unlike most other forms of mental handicap, it occurs in large numbers, it is of known genetic origin and, most importantly, it can be confidently identified at birth. Careful, detailed study of early learning ability in children with DS may thus present us with unique insights into developmental processes in handicap.

One of the major debates in research on development in mentally handicapped children is whether developmental processes are broadly equivalent to those seen in normal development (only taking longer to unfold and progressing to a lower ceiling) or whether there are significant differences in both structure and organisation — the "difference versus delay" debate (Zigler and Balla, 1982). Recently, a growing body of evidence has suggested that learning in children with mental handicap may indeed differ in very fundamental ways from that seen in non-handicapped children, with important qualitative as well as quantitative differences existing between the two populations (Rogers, 1977; Morss, 1983, 1985; Macpherson, 1984; Rondal, 1984, 1988; Cherkes-Julkowski *et al.*, 1986; Wishart, 1986, 1987; Duffy and Wishart, 1987; Moss and Hogg, 1987; Dyer *et al.*, 1988). If this is the case, both developmental theory and educational practice will require substantial revision.

Despite the increasing evidence that a delay theory may not adequately describe developmental processes in handicap, a surprisingly large number of studies have prejudged the issue in their experimental design, either explicitly or implicitly. This is particularly obvious in studies in which tasks designed for and standardised on young, non-handicapped children are presented to older, handicapped children. Failure is typically interpreted as evidence that a skill has yet to be acquired, that there is still a basic lack of competence in the area under study. Without evidence of the suitability of such tasks for children of that age, whether handicapped or non-handicapped and in the absence of developmental data on young handicapped children on these same tasks, such conclusions cannot be justified. The tasks may simply have failed to engage the older child, a case of "won't do" rather than "can't do" (Koegel and Mentis, 1985; Wishart, 1987).

The assumption of slow development may be less obvious but is equally prevalent in other types of study. The widespread practice of matching handicapped and non-handicapped children on the basis of mental age (MA), for example, also assumes that developmental processes are identical in both sets of children and that both have reached the same stage in development by the same developmental pathways. This assumption seems unwise. It is well known that two children with identical MAs may nonetheless have very different developmental profiles. The validity of MA scores as a measure of cognitive ability, moreover, depends on competence being accurately reflected in performance. Given the substantial and inevitable differences in learning histories, both in terms of length and experience of success/failure, it seems unlikely that performance and competence will be similarly linked in handicapped and non-handicapped children. This suggestion is backed up by findings from a number of studies which indicate that differences in performance between MA-matched handicapped and non-handicapped children are often due to differences in motivation, the handicapped child frequently underperforming in many situations (Snyder, 1977; Siegel, 1979).

Trying to decide between the relative merits of the delay and difference arguments may in fact turn out to be a fruitless pursuit, as impossible a task as trying to portion out the exact and relative contributions of genes and environment to development. To some extent, much of the argument depends on the level of analysis employed: studies in which the evidence appears to support a slow development theory may, on more detailed analysis, reveal the existence of importance differences (Kiernan, 1984). What it does draw our attention to, however, is the need for more detailed information on development in children with handicap, and especially for a better description of the earliest stages in that development. Until recently, surprisingly few data were available in sufficient detail to allow insightful comparison of developmental processes in handicapped and non-handicapped children. The majority of studies simply told us what we already know — that children with mental handicap do less well on almost any psychological test than do their non-handicapped peers.

A growing number of studies are now directed at filling in these gaps in our understanding. The authors are presently conducting a series of interlinked cross-sectional and longitudinal studies of early learning and development in infants and young children with DS. These in-depth studies aim not only to assess the deficits in learning but also to identify the processes which impede that learning. The longer term aim is to identify the learning "style" of the child with DS and from that, to generate effective teaching procedures (for an overview, see Wishart, 1989). In the particular study to be described here, two tests of cognitive ability in infancy were presented on two occasions to a group of 18 infants and young children with DS. The two tests used were the Mental Scale of the Bayley Scales of Infant Development (BSID) and a set of Piagetian object concept tasks (covering Piaget's Stages III-VI). These tests were chosen firstly, because a substantial data-base, including detailed norms and reliability measures, was already available for non-handicapped children (Werner and Bayley, 1966; Bayley, 1969; Uzgiris and Hunt, 1975; Wishart and Bower, 1984, 1985) and secondly, because both tests have been widely used in studies of mental development in handicapped children (for reviews see Gibson, 1978; Carr, 1985; Morss, 1985).

The study had three main aims:

(1) to investigate the stability of performance of infants and young children with DS on standardised tests of cognitive development,

(2) to investigate the relationship between performance and competence in children with DS, and

(3) to investigate the adequacy of describing early cognitive development in DS in terms of normal developmental processes.

These aims seemed important for practical and theoretical reasons. The theoretical issues have already been discussed above. On the more practical level, assessment of cognitive ability plays a far greater role in the life of young children with mental handicap than in that of ordinary children (Morss, 1988). Many decisions, most obviously those concerning educational placement, are based on outcome measures from cognitive tests such as the ones used here. More often than not, the test is given only once and generally by an adult who is unfamiliar with the child; frequently the mother is not present. If such tests do not accurately measure mental ability in handicapped children or do not do so reliably, re-evaluation of the value of such testing would be indicated. If, moreover, the inadequacy of these tests results from basic differences in the ways that children with handicap approach learning, a re-assessment of teaching methods would also be required: traditional teaching methods which are highly effective with ordinary children may be ineffective — or even counter-productive — when applied to mentally-handicapped children.

METHOD

Sample

The Mental Scale of the BSID is designed to test for abilities which normally emerge between birth and 30 months; object concept tasks cover a similar age range, birth to 24 months. Cognitive development in children with DS is generally described as progressing at approximately half the normal rate (see, e.g., Berry *et al.*, 1984). In order therefore to examine the validity and reliability of the two chosen tests over their whole range, 18 DS children in the age range 6 months to 4 years were recruited to the study, three at each of the following ages: 6, 12, 18, 24, 36 and 48 months.

All children were confirmed by karyotyping as having Down's Syndrome (DS): all were standard Trisomy 21, with one also showing a small percentage of XXX cells. There were 13 females and 5 males. Nine children had congenital heart disease of varying degrees of severity (the majority minor), three had diagnosed hearing loss, four had minor visual impairments and one suffered from arthritis; none, however, had been hospitalised for any lengthy period during their development. All children had been volunteered by their parents and came from a wide range of socio-economic backgrounds. Seven were first borns. Number of siblings at time of testing ranged between none and four.

None of the children was naive to assessment situations. All were being regularly visited at home by the local authority's home teaching service and all but the three youngest and one 2-year-old had already participated in studies carried out by the present investigators. This is clearly a possible source of confounding variance but is an unavoidable feature of carrying out a research programme with any clinical population of restricted size and with this population in particular. As the average interval since last being seen by the authors was three months, it was considered unlikely that this factor would seriously affect the validity of this study.

Each child in the present study served as his or her own control. Since both tests had already been standardised on large groups of non-handicapped children and since the specific purpose of the study was to investigate the validity and reliability of the two chosen tests within this particular population, no direct control group of non-DS children was considered necessary. As indicated in the introduction, the present authors have in any case severe reservations, both practical and theoretical, about the value of using control groups matched on the basis of mental age. It was hoped that the present study would provide additional empirical support for this viewpoint. Controls matched for chronological age also seemed to have little to recommend them, being likely only to confirm what is already known — that a child with DS will perform less well on cognitive tests than a child of the same age who has no handicap (Duffy and Wishart, 1987; Wishart, 1987).

Procedure

All testing took place in the Department of Psychology in a small, quiet room with minimal distracting features. Testing was scheduled for a time of day when the mother thought her child was most likely to be alert and co-operative. The mother was present throughout and the child sat on her knee for testing; mothers were asked to encourage their children if necessary but not to direct their responses in any way.

Two sets of tasks, widely used to assess early cognitive ability, were presented: a series of Piagetian object concept tasks (presented by the first author), followed by items from the Mental Scale of the BSID. Testing was repeated one to two weeks later, with procedure identical in both sessions. All testing sessions were videorecorded for subsequent analysis.

Object concept tasks: Object concept development is indexed by performance on a series of search tasks in which the hiding sequence increases in complexity according to the stage being tested (Piaget, 1937). In all but the simplest task, the infant must choose between one of two identical occluders in order to retrieve a hidden toy; unless he or she fully comprehends the particular hiding sequence, performance can only be at chance level.

Four levels of task, covering Stages III-VI of object concept development, were presented in ascending order of difficulty: Task 1, a simple 1-occluder hiding task; Task 2, a task involving retrieving an object from inside one of two possible occluders, the object having been visibly displaced from one to the other (widely known as the "AAB" task); Task 3, an inference task, a task in which, on discovering the absence of the object from an expected location, the child must infer its true location; and Task 4, a 2-occluder "switching" task in which, following the hiding of the object, the positions of the two occluders are transposed before the child is allowed to search. White cardboard cups served as occluders and a variety of small toys as objects. Four trials of each task were given, with side of hiding randomised over trials. Success on all four trials was required for accreditation of a pass on any given level of task.

Exact details of task procedure and a full justification of scoring criteria adopted can be found in Wishart and Bower (1984). In that study, the ages by which 75 per cent of infants could pass Tasks 1 to 4 were 5, 10, 15 and 22+ months respectively.

Mental Scale of Bayley Scales of Infant Development (BSID): Items from the BSID Mental Scale were presented in accordance with the procedural instructions given in the test manual (Bayley, 1969). A basal level was established and higher level test items then presented until a minimum of six consecutive failures had been recorded. From the raw score thus obtained, children were given an age equivalent (AE) level and, where possible, an MDI (Mental Development Index) score from the BSID normative tables.

RESULTS

Both qualitative and quantitative analyses of performance on the two task sets were carried out.

Object concept tasks

Success rates: Scores from the two testing sessions for each of the four levels of object concept task are shown in Table 1.

Success rates were not impressive at any age in either session. If total scores are considered, there was some evidence of scores increasing with increasing age (Jonckheere $S = 59$; $P<0.05$); this increase was very modest, however, given the norms for these tasks and the large difference in age between youngest and oldest children ($3^1/_2$ years). With the exception of Task 1, success rates were generally poor on all levels of task. The results from Task 2 were surprising. The age by which most non-handicapped children can pass this task is 10 months. As might have been expected, the majority of 6-12 month-old handicapped children failed in both sessions; all but one of the 18-24 month-olds passed both times. This relatively easy task appeared to present difficulties to the 3-4 year-olds, however, with four failing in at least one session. Task 3 showed a more usual developmental pattern, with young children failing and only the very oldest children succeeding. Task 4 seemed difficult at all ages, although there was some improvement in older children by the second testing session.

TABLE 1

PASS RATES ON EACH OF THE FOUR LEVELS OF OBJECT CONCEPT TASKS IN THE TWO TESTING SESSIONS

	Task 1			Task 2			Task 3			Task 4				
Testing Session	I	II		I	II		I	II		I	II			
Age (months)													Pass rate (all tasks combined)	N Child varying over sessions
6	1	2	(1/0)	0	0	(0/0)	0	0	(0/0)	0	0	(0/0)	3	1
12	3	3	(0/0)	2	1	(1/2)	0	0	(0/0)	0	1	(1/0)	10	4
18	3	3	(0/0)	3	3	(0/0)	0	0	(0/0)	0	1	(1/0)	13	1
24	3	3	(0/0)	3	2	(0/1)	0	0	(0/0)	0	0	(0/0)	11	1
36	3	3	(0/0)	1	2	(1/0)	3	3	(0/0)	0	1	(1/0)	16	2
48	3	3	(0/0)	1	1	(0/0)	3	2	(0/1)	0	2	(2/0)	15	3
Totals:	16	17	(1/0)	10	9	(2/3)	6	5	(0/1)	0	5	(5/0)	68 (8/4)	12

Note: Numbers in parentheses: N children whose performance improved over sessions/ N children whose performance deteriorated.

Stability of pass/fails: Although the grouped data indicated that success rates varied very little between sessions, this was not the case when individual records were examined. The performance of 12/18 children differed on at least one task in the two testing sessions (see Table 1). In eight cases, performance improved while four children failed to repeat an earlier success on re-testing (one 12-month-old improved performance on Task 4 but failed at a previously-passed level, on Task 2).

Order of difficulty of tasks: Normative studies have established that these tasks form a developmental hierarchy, ascending in order of difficulty from Task 1 through to Task 4: a child who can pass Task 3, for example, can reliably be expected to pass Tasks 1 and 2 but not Task 4. Group scores for each of the four levels of task appeared to confirm this as the order of difficulty pertaining in the DS children tested here (see Table 1). Closer examination of individual patterns of pass/fails did not support this conclusion, however.

Differences in difficulty level between any two tasks on either testing session were computed for each child and the derived difficulty order compared with the normatively-derived sequence. As in any task order analysis of cross-sectional data, each child provided data on the relative difficulty of only a subset of the tasks presented, those tasks falling to either side of his or her current developmental stage; any child who either failed or passed all tasks (two in this study) provided no relevant data. (For example: a child who passed Tasks 1, 2 and 4 but failed Task 3, provided no data on the relative difficulty of Tasks 1, 2 and 4 to each other but did provide data on the difficulty of Task 3 in relation to each of these tasks; since, moreover, Task 3 was found to be more difficult than Task 4, the derived difficulty order in this case could be classed as deviating from the normatively-derived hierarchy for these tasks.)

In six out of the 16 cases where comparative information on task difficulty existed, order of difficulty did not match the norm for at least one task pairing ($\chi^2 = 6.58$ (df1): $P<0.01$). In two younger children, this was due to a reversal in difficulty level between Tasks 3 and 4; in four older children, it was due to Task 2 proving more difficult than Task 3.

Failures to engage: Over the two testing sessions, there were 76 instances where a child's performance did not meet the stringent statistical criteria necessary to score a pass. Inspection of performance protocols, however, suggested that many of these failures could not confidently be attributed to a straightforward absence of the required level of cognitive ability. In 55 cases, there was evidence that the child had not fully engaged in all of the four trials presented (Table 2), a frequency which would suggest extreme caution in using such data as a basis for assessing cognitive status in these children.

TABLE 2

FREQUENCIES OF FAILURES TO ENGAGE ON EACH OF THE FOUR OBJECT CONCEPT TASKS IN EACH AGE GROUP

Age (months)	Session No.	1 t/p	2 t/p	3 t/p	4 t/p	Totals t/p
6	I	1/0	2/0	2/0	2/0	7/0
	II	0/0	2/0	1/1	2/0	5/1
12	I	0/0	1/0	1/2	2/0	4/2
	II	0/0	0/1	0/3	0/1	0/5
18	I	0/0	0/1	1/1	0/0	1/2
	II	0/0	0/0	1/2	0/1	1/3
24	I	0/0	0/0	1/1	1/0	2/1
	II	0/0	1/0	1/1	1/0	3/1
36	I	0/0	1/2	0/0	3/0	4/2
	II	0/0	0/2	0/0	1/2	1/4
48	I	0/0	1/1	0/0	0/1	1/2
	II	0/0	1/1	1/0	1/0	3/1
Totals:		1/0	9/8	9/11	13/5	32/24

Note: t = total failure to engage; p = partial failure to engage (see text for definitions)

Failure to engage occurred with near equal frequency in both sessions. It was defined as any instance where a child refused to attend to the hiding sequence or failed to search quickly and clearly for the hidden object (the success or failure of any search being immaterial). Failure to engage was classified as being either total or partial. It was classified as total if on any trial, the subject either (a) refused to attend to the full hiding sequence or (b) made no attempt at search. Since the criterion for a pass on all tasks was 4/4, this level of "switching out" meant that some children failed by default, rather than by making errors. Interestingly, many such children would happily re-engage, albeit it often only temporarily, in subsequent trials of another task, even although that task involved exactly the same test materials and was of a higher level of difficulty than that just refused. Also noteworthy was that all but the very youngest children were as likely to opt out after a trial in which they had successfully retrieved the object as after one in which initial search had been unsuccessful.

Failures to engage were often only partial. Although prepared to watch the hiding sequence and to search, it was clear that that search did not represent the child's best attempt at the problem being presented. As trials progressed, children would delay search, often attempting to divert the tester into some sort of social game by producing some sort of party-trick (such as clapping hands) or response which was quite unrelated to the task in hand. This sort of behaviour can only make a difficult task more difficult, adding a memory component where none previously existed. Some children simply opted out of a particular task by refusing to commit themselves to either cup, sweeping both from the table, an inefficient, low-level strategy which led to recovery of the object but did not meet the criteria for a pass.

Sided strategies: Less readily classifiable as refusals to engage (and not included in Table 2) were instances where children adopted a sided strategy, automatically choosing the cup on one side on every trial of a given task, irrespective of whether that cup had actually been involved in the relevant hiding sequence or not. There were 14 instances of a sided strategy being used, all but one occurring in relation to the two higher level tasks, Tasks 3 and 4.

These rates do not differ significantly from the frequencies of usage that would be expected on the basis of normative studies of these tasks. However, the quality of the behaviour and the response to its outcome were noticeably different to the first author, a highly experienced tester both of these tasks and of this population. A sided strategy is "correct" as often as it is "wrong" in the design adopted here.

Typically, however, on trials where the strategy in fact failed, children showed no surprise or concern, simply going on quickly to search at the remaining cup. The side repeatedly chosen, moreover, corresponded to the child's demonstrated hand preference in all but two instances. This is not unusual for Task 4 but is highly unusual for Task 3, where a sided strategy is generally the result of a tendency to repeat search to the side which first produced success. This suggests that while on the surface these children appeared to be co-operating, there was little true task engagement, the strategy adopted being one of convenience rather than one evolved from cognitive considerations.

A number of other indications of falling levels of task engagement were repeatedly observed within trial sets of a given task. Attention became increasingly difficult to capture and often drifted during hiding sequences, and search was clearly less motivated. In Task 3, for example, exhaustive search of all possible locations in early trials fell to 1-stage search in later trials: both qualify as "fails" on statistical criteria but the number of locations searched in this task has been shown in normative research to be a good indicator of developmental status. Similar evidence of reduced effort was found for Task 4; errors — quickly corrected in initial trials — were simply accepted in later trials, the child sitting, as if helpless, waiting to be shown where the toy was. Since introduction of a new level of task led to a revival of effort and attention, this lower-level responding could not be attributed to simple fatigue effects, to basic attentional deficits or to an overall low level of motivation.

Mental Scale of Bayley Scales of Infant Development (BSID)

Raw scores: In over half the testing sessions, raw scores on the BSID Mental Scale were too low to be converted into Mental Development Indices (MDIs). BSID data were therefore treated in terms of raw scores and their age-equivalent (AE) levels.

Table 3 shows the mean raw scores in the two testing sessions for the six age groups tested. Raw scores increased significantly with increasing age (Jonckheere $S = 127$; $P<0.01$ for both sessions). Even the six-month-olds, however, scored 1 to 2 months below the norm for their age, a deficit in performance which widened with increasing age, the oldest children scoring at levels between 18 and 28 months behind expected levels for non-handicapped children of similar age.

TABLE 3
RAW SCORES AND VARIABILITY ON THE BSID MENTAL SCALE IN THE SIX AGE GROUPS

Age (months)	Testing Session I		Testing Session II		Total N items varying over sessions	
	Mean score (AE)				Fail-to-pass	Pass-to-fail
6	55.3	(4.5)	53.3	(4.5)	12	19
12	77.3	(7.0)	75.6	(7.0)	10	15
18	100.0	(11.0)	99.7	(11.0)	11	13
24	109.0	(13.0)	108.7	(13.0)	16	16
36	131.3	(19.0)	130.3	(19.0)	12	15
48	142.3	(22.0)	142.3	(22.0)	13	13
Totals:					74	91

Note: AE = age equivalent level (in months)

Stability of pass/fails: In order to examine the stability of pass/fails, two measures of reliability were taken: a comparison of the raw scores achieved by each child in the two testing sessions (the most direct measure of the reliability of an "intelligence" test) and a more detailed analysis of the item-item consistency for each child over the two testing sessions. Quite different pictures emerged from these two analyses. Comparison of raw scores indicated that there was no significant difference in performance over the two sessions ($t = 0.76$, (df 17); NS). Examination of individual performance profiles, however, showed that even where raw scores were very similar (or even identical), performance had in fact changed on a substantial number of items (see Table 3).

Item-item agreement over sessions in individual children would be the most obvious and sensitive measure of a test's reliability. No such normative data are provided by Bayley, however. Re-test data from a very small sample of children (28) at only one age level (8 months) are presented and no indication of the reliability of test items for individual children is given, only the relative reliability of individual test items. Procedure, moreover, varied from the standard testing procedure in that all re-test children were given the same number of items, irrespective of their individual ability levels. In order, therefore, to investigate whether variability in performance was greater in the sample of DS children tested here than in the normative sample, the variability in each individual child's scores over the two testing sessions was compared to the variability which would be predicted on the basis of the standard errors of measurement (SEm) provided in the BSID manual for the range of mental ages covered by the test. As a measure of reliability, this is likely to be an over- rather than under-estimate of the variability in scores to be expected in individual children.

In most cases, scores were too low to be converted into MDIs. Comparisons were therefore carried out against the variability predicted in raw scores on the basis of the SEm and AE level for each child (for a child of 24 months with an age equivalent level of 12 months, for example, raw scores would be expected to vary by a maximum of 6 points on re-testing). A t-test indicated that there was a significant difference between observed and expected variability over the two sessions (t = 2.78, df (17), P<0.01). Overall, there were 165 cases of item instability: in 74 cases, items which had been failed in the first session were passed in the second session; in 91 cases, the change was in the opposite direction, a previously passed item being failed on re-testing ($\chi^2 = 0.96$ (df 1): NS). Degree of variability and ratio of pass-to-fail to fail-to-pass changes were similar at all ages (t = -0.82, (df17); NS).

Patterns of fails-to-passes and of passes-to-fails are equally interesting but for different reasons. A fail-to-pass could result from a genuine developmental advance having been made between sessions. However, given the short interval between testing sessions, this was unlikely to be the case when performance improved on a substantial number of items in the second session. It was also unlikely that any such improvement could have resulted merely from increased familiarity with the tester on the second session, since she was already familiar to most of the children on the first testing session. More likely is that fluctuating levels of task engagement were responsible for the instability shown. In this respect, a pass-to-fail pattern is far less ambiguous. Clearly the required behaviour *was* in the child's repertoire but was not for some reason being reproduced on the second testing. This instability must inevitably lead to underestimation of the child's true competence. Unlike fails, any pass, whether produced in the first or the second session, is likely to be genuine; few of the BSID items permit chance success. When the children tested here were credited with all items on which they had scored a pass, regardless of its stability, this "optimal" score proved — not surprisingly — to be consistently higher than the score actually achieved in either of the two testing sessions (1st/2nd session: t = 2.60/3.24 (df17); P<0.01/0.005).

A further analysis of the difficulty level of those items which proved to be unstable was carried out in which each item which varied was related to the developmental level of each child. The midpoint between lowest basal and highest ceiling levels attained over the two sessions was determined for each child and the numbers of pass-to-fail and fail-to-pass items which fell below and above this midpoint were calculated. Numbers of fail-to-pass items fell equally on either side of this midpoint. There was, however, a trend for more pass-to-fail items to occur above than below this midline and for these to exceed the number of fail-to-pass items above this point (t = 1.40/1.24 (df17); P<0.10/NS respectively).

Failures to engage: Given how unstable performance was, it seemed worth investigating the possible source of this instability. Failures to engage were more difficult to define in relation to the BSID items. Unlike the object concept tasks, BSID items vary greatly, both in content and in terms of task demands. It is particularly difficult to identify with any confidence any failure to engage in relation to items at the bottom end of the scale, where to be credited with a pass the child has simply to respond to a stimulus (e.g., regards cube, turns head to rattle). In these cases, it is impossible to distinguish between a criterion fail, indicating absence of the required level of cognitive ability, and a fail by default, where failure is due to the child having refused to engage. At all levels, the BSID also includes a number of "incidental" items, items in which there is no particular task to be engaged in (e.g., vocalises once or twice, repeats performance laughed at); by definition, these too cannot be incorporated into any failure-to-engage analysis. Forty-seven out of the 165 cases of item instability were omitted from the analysis below on these grounds, although it is perhaps worth noting that as performance did differ on these items (and not in all cases in the fail-to-pass direction), it seems highly probable that something other than lack of the required cognitive ability was determining poor performance in one or other session.

In the 118 remaining eligible cases, subject records were re-examined for any evidence of failure to engage on the failed item. Failure to engage was defined as any instance where the child's response fell into one of the following three categories: (1) refusal to attend to the task, (2) rejection of the task object by casting, swiping or dropping it on the floor, or (3) repeated production of an inappropriate, off-task behaviour in response to the task materials. In over half of the cases where performance on items differed over sessions, failure was clearly due to the child refusing to engage in the task: in 40/64 cases in the pass-to-fail direction, and in 26/54 in the fail-to-pass direction.

DISCUSSION

The low levels of ability demonstrated were somewhat discouraging, given the effort and funds presently being diverted into early intervention programmes for DS children. Unlike DS children in the 1960s, all of the children in this study had home teachers, all were being brought up in their own homes and all had highly motivated parents, all of whom believed in actively encouraging their child's development in all areas. Ability levels differed very little, however, from those seen in earlier studies of this population (for overview, see Gibson, 1978). On average, performance was at approximately half the level that would have been expected given their chronological ages, a finding consistent with recent results from studies in the USA, Australia, Canada and Germany (Pueschel, 1984; Rauh *et al.*, 1987).

Although at one level the results lend support to a "slow development" theory, there did seem to be important differences in the way in which the children with DS responded to both sets of tasks. These particular tasks were selected for study because although neither test was specifically designed for or standardised on handicapped children, both have been used widely to assess early cognitive development in handicapped children. Object concept development was considered by Piaget to be prototypical of all cognitive development, and fifty years on object concept studies still dominate much of the literature on learning in handicapped and non-handicapped children (for reviews, see Schuberth, 1982; Harris, 1988). In a review of psychological tests of early cognitive ability, Francis *et al.* (1987) described the Bayley Scales as "unrivalled at determining a child's developmental status in relation to its age-mates". Only a few studies suggest that either test has any great predictive validity (e.g., Wachs, 1975; Siegel, 1981), but it is generally accepted that each has both concurrent and construct validity — that is, that they reliably measure in a single testing session whether a child has or has not reached important stages in early cognitive development.

The present results would suggest that neither test does this in the case of mentally handicapped children. On both sets of tasks, the variability in performance was far greater than would have been expected on the basis of large-scale normative studies (Bayley, 1969; Horner, 1980; Morss, 1985; Wishart and Bower, 1985). In the object concept tasks, there was no simple pattern to the variability demonstrated, either in terms of age or task level. The performance of 12 children varied over sessions and all levels of tasks showed some instability (although it was most frequently seen in Tasks 2 and 4). "Switching out", however, was consistently associated with *all* of these apparent reversals in ability, children failing to complete all four trials or responding in an inappropriate or stereotyped fashion in either the first or second testing session. Failures to engage were equally prevalent in response to the Bayley items and it was clear from the analysis of "optimal" versus typical performance that in these tasks too, virtually all the children were underperforming to some degree in both sessions.

The many failures to engage suggest that motivational factors play an important role in the expression of competence in performance in young children with DS. In the object concept tasks, this "switching out" may have been responsible for the disruption found to the order of difficulty these tasks normally hold to each other. Longitudinal research with 3-5 year-olds and with infants with DS has indicated that failure on object concept tasks at later ages may not, as is commonly concluded, represent a failure to have acquired the requisite cognitive skills but may result, rather, from a failure to consolidate these skills when they were first acquired, in early infancy (Wishart, 1987; 1988a). In these studies, failure to engage appeared to result from the operation of two distinct processes, both of them linked to the relationship between the difficulty level of the task being presented and the child's stage in development at time of testing. Infants in these longitudinal studies went to considerable lengths to avoid tasks of a level of difficulty one step or more above their current developmental status. This was perhaps not surprising but unfortunately tasks previously mastered also fell outwith the range of engagement, performance typically falling away in the months following initial acquisition. Only tasks falling within a narrowly defined range of difficulty were fully engaged in and reliably passed, suggesting that this form of cognitive avoidance adversely affects both the acquisition and consolidation stages of learning. The analysis of the Bayley data lends some support to this interpretation, with "switching out" and instability occurring in relation to items at both the easy and difficult ends of the child's developmental range.

In both these previous Wishart studies and in the study reported here many task failures were by default, the result of the child refusing to participate in all trials of the task being presented. Often the child used — or more accurately, *mis*-used — social skills to get out of the task in hand (Wishart, 1988b). To classify such responses as "fails" or to describe the child as "untestable" is not particularly helpful, either to the child or to our understanding of that child. It is clear, however, that any attempt to modify performance in such situations will require a better understanding of developmental processes in children with handicap than that presently held. We shall also need deeper insight into the interrelationship of motivational factors, competence and performance in such children. In the case of some failures, it may be possible to modify testing procedures to suit the child's cognitive "style" better; it may be that the number of trials dictated by statistical considerations exceeds that considered to be cognitively or motivationally relevant to the post-acquisition child with DS. Overcoming poor engagement in the acquisition stage may be more difficult but it is perhaps possible that artificial enhancement of success/failure ratios could prevent avoidance behaviour becoming the routine response to difficult learning situations (Seligman, 1975).

Obviously, a cross-sectional study such as this cannot provide any direct evidence on development or developmental processes and cannot therefore explain the mismatch between performance and competence found in so many of the children here. The instability shown could result from poor motivation, could reflect a basic instability in the learning process itself or could be due to a complex developmental interaction between inadequate motivation and inefficient learning processes, with poor motivation contributing directly to the growth of deficits in cognitive functioning. What can safely be concluded is that, if the instability demonstrated in assessment situations is characteristic of the DS child's response to his or her everyday learning environment, quite clearly inefficient use is being made of whatever level of ability is present, a factor which can only add to the child's already existing handicap.

Unstable performance may underlie a feature frequently commented on in relation to development in DS. Parents of DS children often comment that their child's development seems to be very uneven, several advances being made in a comparatively short period of time, only to be followed by a lengthy period in which development seems to have come to a virtual standstill. A number of researchers have found evidence supporting the suggestion that development in DS shows such "plateaux", periods which they have interpreted as periods of consolidation of recently acquired skills (e.g., Carr, 1975; Cunningham, 1988). Others are less optimistic, seeing these plateaux as evidence of the child meeting a "developmental wall", a period during which no development takes place at all (Gibson, 1978). Neither may in fact be an accurate description of what is going on. If individual profiles are monitored, we may find that these apparent lulls in development are as much due to the dropping out of previously passed items as to any failure to make progress on new, higher level tasks.

A longitudinal study presently in progress aims to test out this alternate hypothesis. If this does show poor stability of new acquisitions, clearly teaching effort will have to be directed not only at encouraging development of new skills but also at ensuring that recently acquired skills are adequately consolidated. How easy it will be to bypass the DS child's apparent inclination to avoid this remains to be seen. Some evidence does exist that tailoring of teaching methods to the particular needs and skills of the handicapped child can improve performance and can lead to positive carry over effects in less structured learning situations (see, e.g., Duffy and Wishart, 1987); it is not yet clear, however, whether any long-term benefits accrue from this type of approach (Wishart, 1986). Amplified developmental instability may, as Shapiro (1983) suggests, be built into the learning system, an inevitable effect of the trisomy on development, affecting all but the most buffered areas of development; the degree to which this instability may be compensated for through environmentally-based intervention remains an open question.

ACKNOWLEDGMENTS.—This research was supported by MRC Project No. G8703875N and by ESRC Linked Studentship Award No. C00428725008. We would like to acknowledge, with very many thanks, the help of all of the mothers and infants who took part in this study. We are also grateful to the Scottish Down's Syndrome Association for the constant encouragement and support which they have given to us throughout our research programme. Thanks are also extended to the Department of Child Life and Health, University of Edinburgh, to the Cytogenetics Lab, Royal Infirmary for Sick Children (Edinburgh) and to the Department of Human Genetics, Western General Hospital (Edinburgh) for providing karyotype information.

Correspondence and requests for reprints should be addressed to Dr Jennifer Wishart, Edinburgh Centre for Research in Child Development, Department of Psychology, 7 George Square, Edinburgh EH8 9JZ, Scotland.

REFERENCES

BAYLEY, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corporation.

BERRY, P., GUNN, V. P., AND ANDREWS, R. J. (1984). Development of Down's Syndrome children from birth to five years. In BERG, J. M. (Ed.), *Perspectives and Progress in Mental Retardation*. Baltimore: University Park Press.

CARR, J. (1975). *Young Children with Down's Syndrome*. London: Butterworth.

CARR, J. (1985). The development of intelligence. In LANE, D., and STRATFORD, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

CHERKES-JULKOWSKI, M., GERTNER, R., AND NORLANDER, K. (1986). Differences in cognitive processes among handicapped and average children: a group learning approach. *J. Learn. Disabil.*, 19, 438-455.

CUNNINGHAM, C. (1988). *Down's Syndrome: An Introduction for Parents. (2nd ed.)*. London: Souvenir Press.

DUFFY, L., AND WISHART, J. G. (1987). A comparison of two procedures for teaching discrimination to Down's Syndrome and normal children. *Br. J. educ. Psychol.*, 57, 265-278.

DYER, S., GUNN, P., RAUH, H., AND BERRY, P. (1988). Motor development in Down Syndrome children: an analysis of the Motor Scale of the Bayley Scales of Infant Development. Paper presented at the 8th IASSMD Congress, Dublin, Ireland, August.

FRANCIS, P. L., SELF, P. A., AND HOROWITZ, F. D. (1987). The behavioural assessment of the neonate: an overview. In OSOFSKY, J. D. (Ed.), *Handbook of Infant Development (2nd ed.)*. New York: Wiley.

GIBSON, D. (1978). *Down's Syndrome: The Psychology of Mongolism*. Cambridge: Cambridge University Press.

HARRIS, P. L. (1988). The development of search. In SALAPATEK, P., and COHEN, L. B. (Eds.), *Handbook of Infant Perception*. New York: Academic Press.

HORNER, T. M. (1980). Test-retest and home-clinic characteristics of the Bayley Scales of Infant Development in nine- and fifteen-month-old infants. *Child Dev.*, 52, 751-758.

KIERNAN, C. C. (1984). The development of communication and cognition. In DOBBING, J. *et al.* (Eds.), *Scientific Studies in Mental Retardation*, London: Macmillan Press.

KOEGEL, R. L., AND MENTIS, M. (1985). Motivation in childhood autism: can they or won't they? *J. Child Psychol. Psychiat.*, 26, 185-191.

MACPHERSON, F. D. (1984). The structure of sensori-motor intelligence in severely and profoundly mentally handicapped children. Unpublished PhD thesis, University of Southampton, England.

MORSS, J. R. (1983). Cognitive development in the Down's Syndrome infant: slow or different? *Br. J. educ. Psychol.*, 53, 40-47.

MORSS, J. R. (1985). Early cognitive development: difference or delay? In LANE, D., and STRATFORD, B. (Eds.) *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

MORSS, J. R. (1988). Cognitive motivation in the young child with Down's Syndrome. Paper presented at the 8th IASSMD Congress, Dublin, Ireland, August.

MOSS, S., AND HOGG, J. (1987). The integration of manipulative movements in children with Down's Syndrome and their non-handicapped peers. *Human Movement Science*, 6, 67-99.

PIAGET, J. (1955). *The Construction of Reality*. London: Routledge and Kegan Paul. (Original French edition, 1937).

PUESCHEL, S. M. (1984). *The Young Child with Down Syndrome*. New York: Human Sciences Press.

RAUH, H., BERRY, P., AND RUDINGER, G. (1987). Patterns of development in Down Syndrome children using the Bayley Scales. Paper presented at the Ninth Biennial Meeting of ISSBD, Tokyo, Japan, July.

ROGERS, S. J. (1977). Characteristics of the cognitive development of profoundly retarded children. *Child Dev.*, 48, 837-843.

RONDAL, J. A. (1984). Linguistic and pre-linguistic development in moderate and severe mental retardation. In DOBBING, J. *et al.* (Eds.) *Scientific Studies in Mental Retardation*. London: Macmillan Press.

RONDAL, J. A. (1988). Language development in Down's Syndrome: a life-span perspective. *Int. J. behav. Dev.*, 11, 21-36.

SCHUBERTH, R. E. (1982). The infant's search for objects: alternatives to Piaget's theory of object concept development. In LIPSITT, L. P., and ROVEE-COLLIER, C. K. (Eds.), *Advances in Infancy Research: Vol. 2*. Norwood, N. J.: Ablex.

SELIGMAN, M. E. P. (1975). *Helplessness*. San Fransisco: W. H. Freeman.

SIEGEL, P. S. (1979). Incentive motivation and the mentally retarded person. In ELLIS, N. P. (Ed.), *Handbook of Mental Deficiency: Psychological Research and Theory*. Hillsdale, N.J.: Erlbaum.

SIEGEL, L. (1981). Infant tests as predictors of cognitive and language development. *Child Dev.*, 52, 545-557.

SHAPIRO, B. L. (1983). Down Syndrome — a disruption of homeostasis. *Am. J. med. Genet.*, 14, 241-269.

SNYDER, L. S. (1977). Communicative and cognitive abilities and disabilities in the sensori-motor period. *Merrill Palmer Q.*, 24, 161-180.

UZGIRIS, I., AND HUNT, J. MCV. (1975). *Assessment in Infancy: Ordinal Scales of Psychological Development*. Urbana, Ill.: University Press.

WACHS, T. D. (1975). Relation of infants' performance on Piaget scales between 12 and 24 months and their Stanford-Binet performance at 31 months. *Child Dev.*, 46, 929-935.

WERNER, E. E., AND BAYLEY, N. (1966). The reliability of Bayley's Revised Scale of Mental and Motor Development during the first year of life. *Child Dev.*, 37, 39-50.

WISHART, J. G. (1986). The effects of step-by-step training on cognitive performance in infants with Down's Syndrome. *J. ment. Defic. Res.*, 30, 233-250.

WISHART, J. G. (1987). Performance of young non-retarded children and children with Down Syndrome on Piagetian infant search tasks. *Am. J. ment. Defic.*, 92, 169-177.

WISHART, J. G. (1988a). Early learning in infants and young children with Down's Syndrome. In NADEL, L. (Ed.), *The Psychobiology of Down Syndrome*. Boston: M.I.T. Press.

WISHART, J. G. (1988b). Misuse of social skills in young children with Down's Syndrome. Paper presented at the summer meeting of the Experimental Psychology Society, Edinburgh, Scotland, July.

WISHART, J. G. (1989). Learning to learn: the difficulties faced by infants and young children with Down's Syndrome. In FRASER, W. I. (Ed.), *Key Issues in Research in Mental Retardation*. London: Routledge (in press).

WISHART, J. G., AND BOWER, T. G. R. (1984). Spatial relations and the object concept: a normative study. In LIPSITT, L. P., and ROVEE-COLLIER, C. K. (Eds.), *Advances in Infancy Research: Vol. III*. Norwood, N.J.: Ablex.

WISHART, J. G., AND BOWER, T. G. R. (1985). A longitudinal study of the development of the object concept. *Br. J. dev. Psychol.*, 3, 243-258.

ZIGLER, E., AND BALLA, D. (Eds.) (1982). *Mental Retardation: The Developmental Difference Controversy*. Hillsdale, N.J.: Erlbaum.

(*Manuscript received* 3rd *March*, 1989)